

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 120.153 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DB=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US09821883/runat 09092004.105128.6657/app_query.fasta_1.3100
-DB=Issued Patents_NA -QPMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=Dts -START=1 -END=-1 -MATRIX=Dlosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CIGN 1.1 215 @runat 09092004.105128.6657 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2439.5	64.5	4530	1	US-08-229-515A-9
2	2439.5	64.5	4530	1	US-08-645-865-9
3	2439.5	64.5	4530	4	US-09-167-322-4
4	2439.5	64.5	4530	4	US-09-527-487-1
5	2439.5	64.5	4530	4	US-09-877-177A-11
6	2439.5	64.4	4473	2	US-09-048-804-1
7	2434.5	64.4	4473	3	US-09-056-105-26
8	2434.5	64.4	4473	4	US-09-663-834A-3
9	2434	64.3	9274	4	US-09-811-115-1
10	2422	64.0	3768	4	US-09-811-115-2
11	2417	63.9	3768	2	US-08-625-101-1
12	2417	63.9	3768	2	US-08-356-786-1

13	2033.5	53.8	2385	2	US-09-146-283-3	Sequence 3, Appli
14	2033.5	53.8	2385	3	US-08-579-823A-3	Sequence 3, Appli
15	2033.5	53.8	2385	3	US-09-344-195-3	Sequence 3, Appli
16	1976	52.2	3955	1	US-08-229-515A-14	Sequence 14, Appl
17	1976	52.2	3955	1	US-08-645-865-14	Sequence 14, Appl
18	1626.5	43.0	1872	3	US-08-422-108-2	Sequence 2, Appli
19	1626.5	43.0	1872	4	US-08-422-734-2	Sequence 2, Appli
20	732	19.3	1868	1	US-08-658-883B-1	Sequence 1, Appli
21	732	19.3	1868	4	US-09-676-610B-26	Sequence 26, Appl
22	731.5	19.3	1593	4	US-09-676-610B-25	Sequence 25, Appl
23	731.5	19.3	5532	2	US-08-475-035-3	Sequence 3, Appli
24	731.5	19.3	5532	2	US-09-676-610B-17	Sequence 17, Appl
25	723.5	19.1	5484	3	US-09-632-580A-3	Sequence 3, Appli
26	723.5	19.1	5501	1	US-08-484-438-1	Sequence 3, Appli
27	723.5	19.1	5555	1	US-08-484-438-3	Sequence 3, Appli
28	711	18.8	4905	1	US-07-978-895-3	Sequence 3, Appli
29	711	18.8	4905	1	US-08-473-119-3	Sequence 3, Appli
30	711	18.8	4905	2	US-08-475-352-3	Sequence 3, Appli
31	711	18.8	4905	4	US-09-170-699-3	Sequence 3, Appli
32	711	18.8	4975	3	US-09-630-706-3	Sequence 3, Appli
33	703	18.6	1958	4	US-09-570-454-1	Sequence 1, Appli
34	703	18.6	1958	4	US-09-867-521-1	Sequence 1, Appli
35	624.5	16.5	4545	6	5183884-3	Patent No. 5183884
36	533.5	14.1	1588	2	US-09-146-283-1	Sequence 1, Appli
37	533.5	14.1	1588	3	US-08-579-823A-1	Sequence 1, Appli
38	533.5	14.1	1588	3	US-09-344-195-1	Sequence 1, Appli
39	493	13.0	322	1	US-08-421-356-1	Sequence 1, Appli
40	493	13.0	322	4	US-09-046-783-1	Sequence 1, Appli
41	491.5	13.0	1409	4	US-09-634-368-3	Sequence 3, Appli
42	490.5	13.0	896	6	5200327-2	Patent No. 5200327
43	490.5	13.0	900	1	US-08-318-193-7	Sequence 7, Appli
44	487	12.9	905	6	5200327-3	Patent No. 5200327
45	487	12.9	909	1	US-08-318-193-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 4699.88 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLIARAASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO.spool/US09821883/runat_05092004.105127.6643/app.query.fasta_1.3100
-DB=EST -QFWT=fasGap -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THRM SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1 11571 @runat_05092004.105127.6643 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_inv:*
18: em_gss_hum:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rpd:*
26: em_gss_phg:*
27: em_gss_vri:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1970.5	52.1	4323	11	AK031099	Mus muscu AK031099 Mus muscu
2	1960	51.8	4463	11	AK083669	Mus muscu AK083669 Mus muscu
3	1291.5	34.1	1016	12	BM562913	AGENCOURT BM562913 AGENCOURT
4	1169	30.9	1129	12	BM802792	AGENCOURT BM802792 AGENCOURT
5	1167.5	30.9	885	14	CA455074	AGENCOURT CA455074 AGENCOURT
6	1166	30.8	871	14	CA488274	AGENCOURT CA488274 AGENCOURT
7	1162	30.7	4715	11	AF318349	Homo sapi AF318349 Homo sapi
8	1139.5	30.1	893	14	CA455141	AGENCOURT CA455141 AGENCOURT
9	1090	28.8	894	14	CA454131	AGENCOURT CA454131 AGENCOURT
10	1090	28.8	1004	14	CD515356	AGENCOURT CD515356 AGENCOURT
11	1089	28.8	653	12	BM721340	AGENCOURT BM721340 UI-E-E01-
12	1081	28.6	902	14	CA488868	AGENCOURT CA488868 AGENCOURT
13	1073	28.4	588	13	EX478931	DKFZp886A EX478931 DKFZp886A
14	1046	27.7	569	9	AL701765	AL701765 DKFZp886P
15	1033	27.3	583	13	EX479114	DKFZp886P EX479114 DKFZp886P
16	1022.5	27.0	795	14	CA328613	UI-M-PY0- CA328613 UI-M-PY0-
17	1014	26.8	3110	11	AK031542	Mus muscu AK031542 Mus muscu
18	1010	26.7	808	14	CA489534	AGENCOURT CA489534 AGENCOURT
19	991	26.2	616	14	CD721801	oJ01f07.Y CD721801 oJ01f07.Y
20	967	25.6	943	13	BQ958632	AGENCOURT BQ958632 AGENCOURT
21	966	25.5	1988	11	BC023725	Mus muscu BC023725 Mus muscu
22	966	25.5	3372	11	BC046553	Mus muscu BC046553 Mus muscu
23	935	24.7	855	13	BU594980	AGENCOURT BU594980 AGENCOURT
24	924	24.4	813	12	BI557797	603236977 BI557797 603236977
25	921	24.3	614	10	AW370693	QV1-BT026 AW370693 QV1-BT026
26	909.5	24.0	852	14	CD516283	AGENCOURT CD516283 AGENCOURT
27	908	24.0	791	12	BI154872	AGENCOURT BI154872 602902857
28	903	23.9	717	14	CB598701	AGENCOURT CB598701 AGENCOURT
29	901	23.8	491	12	BM790293	K-EST0070 BM790293 K-EST0070
30	897	23.7	906	14	CA454570	AGENCOURT CA454570 AGENCOURT
31	880	23.3	998	12	BI649877	603296516 BI649877 603296516
32	875	23.1	685	9	AU123871	AU123871 AU123871
33	871	23.0	932	14	CA487981	AGENCOURT CA487981 AGENCOURT
34	868	22.9	609	10	AW701942	uq93b02.Y AW701942 uq93b02.Y
35	862.5	22.8	621	12	BG283493	602407782 BG283493 602407782
36	855	22.6	887	13	BQ717097	AGENCOURT BQ717097 AGENCOURT
37	848	22.4	786	12	BI155788	603904360 BI155788 603904360
38	835.5	22.1	649	9	AA496412	AA496412 zv37a02.r
39	817	21.6	533	14	CB437620	685491.MA CB437620 685491.MA
40	817	21.6	552	12	BG277542	ux45d07.Y BG277542 ux45d07.Y
41	817	21.6	670	9	AI906012	RC-BT105- AI906012 RC-BT105-
42	808	21.4	876	13	BQ769889	UI-M-F10- BQ769889 UI-M-F10-
43	805	21.3	514	10	BF998814	QVO-GN014 BF998814 QVO-GN014
44	800	21.1	887	14	CA980253	AGENCOURT CA980253 AGENCOURT
45	792.5	20.9	678	12	BI555157	603236396 BI555157 603236396

ALIGNMENTS

RESULT 1

AK031099

LOCUS

DEFINITION

AK031099

Mus musculus

enriched library, clone:593040N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.

ACCESSION

AK031099

VERSION

AK031099.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

4323 bp mRNA linear HTC 18-SEP-2003
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length enriched library, clone:593040N10 product:v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian), full insert sequence.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi, Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED

3 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koude, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

```

/db xref="FANTOM DB:5930404N10"
/db xref="MGI:2394277"
/db xref="taxon.10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48. -3818
/notes="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
nucleophosmin derived oncogene homolog (avian)
(MGI:MGI:954110, GB|U71126, evidence: BLASTN, 99%,
match=449)"

misc_feature
48. -3818
/notes="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
nucleophosmin derived oncogene homolog (avian)
(MGI:MGI:954110, GB|U71126, evidence: BLASTN, 99%,
match=449)"

ORIGIN
Alignment Scores:
Pred. No.: 1.49e-111 Length: 4323
Score: 1970.50 Matches: 437
Percent Similarity: 36.79% Conservative: 31
Best Local Similarity: 34.36% Mismatches: 73
Query Match: 52.09% Indels: 731
DB: 11 Gaps: 4

US-09-821-883-4 (1-697) x AK031099 (1-4323)
Qy 3 AlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeu 22
Db 33 GCCTGAGTATCATCATGGAGCTGGCGCTGGTGGGGTCTCTCTCGCCCTC 92
Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
Db 93 -----CTGTCCCGCGAGCGCGGGTACC 116
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
Db 117 CAAAGTGTATACCGGTACCGACATGAAGTTCGCACTCCCTGCCAGTCTCTGAGACCCACTG 176
Qy 63 AspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnGlyAsnLeuGluLeuThr 82
Db 177 GACATGCTCCGCCACCTCTACCAAGGCTGTTCAGGTGTGTGAGGGCAATTTGGAGCTTACC 236
Qy 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGlyTyr 102
Db 237 TACTGTCCCGCAATGCCAGCTCTCATTCCTCGAGGACATCCAGGAAGTCCAGGGATAC 296
Qy 103 ValLeuIleAlaHisGlnGlnValArgGlnValProLeuGlnArgLeuArgIleValArg 122
Db 297 ATGCTCATCGCTCACCAACCGAGTGAACACGCTCCCTGCTGAGAGTTTGGCACTGAGAGA 356
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142
Db 357 GGAGCTCAGCTCTTTGAGGACAAGTATGCCCTGGCTGTGTAGACAACCGAGACCCCTTTG 416
Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161
Db 417 GACACGCTCACCCACCGCGCGCCCGCAGGCAAGAACCCAGAGAGGCTGCGGGAGCTGCAGCTT 476
Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuCys 181
Db 477 CGAAGTCTCACAGAGATCTTGAAGGGAGGAGTTTGTATCGTGGGAACCTTCAGCTCTGC 536
Qy 182 TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAlaLeuThr 201
Db 537 TACAGGACATGTTTGTGGAAGGATGCTCCCGTGAAGAATAACAGAGTGGCTCCTCTGTC 596
Qy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
Db 597 GACATGGACACCAATCGTTCCTCCGGGCTGTCCACTTGTGCCCCCAACCTGCAAGACAAT 656
Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrValCysAlaGly 241
Db 657 CACTGTGGGGTGAGAGTCTCTGAAGACTGTCTGAGATCTTGTGTCACCACTCTGACTAGT 716

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 842.741 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARAASLSGLF.....CWKQVKGAPPPPAHHHHH 697

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3304383 seqs, 2515761380 residues
Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app_query.fasta_1.3100
-DB=Published Applications NA -QWMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09821883 @CGN 1.1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

RESULT 1
US-09-821-883-9
; Sequence 9, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graadiss, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct

ALIGNMENTS

1	3783	100.0	2091	9	US-09-821-883-9	Sequence 9, Appl
2	3473.5	91.8	2070	9	US-09-821-883-7	Sequence 7, Appl
3	3014	79.7	1692	9	US-09-821-883-8	Sequence 8, Appl
4	2957.5	78.2	1665	9	US-09-821-883-6	Sequence 6, Appl
5	2478	65.5	1437	9	US-09-821-883-10	Sequence 10, Appl
6	2442.5	64.6	4543	9	US-09-769-508-1	Sequence 1, Appl
7	2439.5	64.5	4530	9	US-09-877-177-11	Sequence 11, Appl
8	2439.5	64.5	4530	13	US-10-392-113-45	Sequence 45, Appl
9	2439.5	64.5	4530	13	US-10-426-836-11	Sequence 11, Appl
10	2439.5	64.5	4530	15	US-10-177-293-125	Sequence 125, Appl
11	2439.5	64.5	4530	15	US-10-007-926A-119	Sequence 119, Appl
12	2439.5	64.5	4530	15	US-10-338-730-1	Sequence 1, Appl
13	2439.5	64.5	4530	15	US-10-101-510-124	Sequence 124, Appl
14	2439.5	64.5	4530	16	US-10-116-275-131	Sequence 131, Appl
15	2439.5	64.5	4530	16	US-10-272-437A-27	Sequence 27, Appl
16	2439.5	64.5	4530	16	US-10-117-937-595	Sequence 595, Appl
17	2439.5	64.5	4530	16	US-10-159-563-208	Sequence 208, Appl
18	2439.5	64.5	4530	16	US-10-435-896-10	Sequence 10, Appl
19	2439.5	64.5	4530	17	US-10-734-564-59	Sequence 59, Appl
20	2439.5	64.5	4642	15	US-10-198-846-10896	Sequence 10896, A
21	2434.5	64.4	4473	10	US-09-441-411-5	Sequence 5, Appl
22	2434.5	64.4	4473	15	US-10-146-473-32	Sequence 32, Appl
23	2434.5	64.4	4473	15	US-10-207-555-44	Sequence 44, Appl
24	2434.5	64.4	4473	15	US-10-101-510-81	Sequence 81, Appl
25	2434	64.3	9274	9	US-09-811-123-7	Sequence 7, Appl
26	2434	64.3	9274	9	US-09-811-115-1	Sequence 1, Appl
27	2422	64.0	3765	15	US-10-207-498-5	Sequence 5, Appl
28	2422	64.0	3768	9	US-09-811-123-8	Sequence 8, Appl
29	2422	64.0	3768	9	US-09-811-115-2	Sequence 2, Appl
30	2422	64.0	3768	13	US-09-984-092-3	Sequence 3, Appl
31	2422	64.0	3768	13	US-10-280-576-3	Sequence 3, Appl
32	2422	64.0	3768	17	US-10-441-779C-3	Sequence 3, Appl
33	2417	63.9	3768	9	US-09-854-356-9	Sequence 9, Appl
34	2417	63.9	3768	9	US-09-930-125-1	Sequence 1, Appl
35	2417	63.9	3768	13	US-09-785-973-1	Sequence 1, Appl
36	2417	63.9	3768	15	US-10-313-644-1	Sequence 1, Appl
37	2414.5	63.8	4606	10	US-09-971-392-70	Sequence 70, Appl
38	1976	52.2	3955	9	US-09-870-759-117	Sequence 117, Appl
39	1976	52.2	3955	9	US-09-854-356-10	Sequence 10, Appl
40	1976	52.2	3955	10	US-09-751-708A-117	Sequence 117, Appl
41	1974.5	52.2	3771	9	US-09-854-356-11	Sequence 11, Appl
42	1644	43.5	2132	16	US-10-412-804A-3	Sequence 3, Appl
43	1641.5	43.4	2149	16	US-10-412-804A-9	Sequence 9, Appl
44	1638.5	43.3	2164	16	US-10-412-804A-5	Sequence 5, Appl
45	1587	42.0	867	9	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-9

Alignment Scores:

Pred. No.: 0 Length: 2091
Score: 3783.00 Matches: 697
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-821-883-4 (1-697) x US-09-821-883-9 (1-2091)

Qy 1 MetArgAlaProLeuLeuAlaArgAlaSerLeuLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGCACCCCTCTCTGGCCAGGAGCAAGCCCTTAGCCCTTCTGTTT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyValaAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAAAGGAGTTGGCGCGGGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCAAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCGCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyValLeuGlu 80
Db 181 CACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCGAGAAACCTGGAA 240
Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACCTACCTGCCACCAATGCCAGCTCTCTCTTCTGCAGGATATCCAGGAGGTGCAG 300
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATCGCTCACAACTGAGGAGGCTCCCACTGCAGAGGCTGCGGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGACCCAGCTCTTTGAGACAACTATGCTTGGCCGTGGCTAGACAAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTTGGGGAGCTGCAG 480
Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTTGAAGCCTCACAGATCTTGAAGGGGGTCTTATCCAGCGAAACCCCGAGCTC 540
Qy 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db 541 TGCTACCAAGGACACGATTTTGGGAAGGACATCTTCCCAAGAAACAACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACATGATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTCTCCGATGTGTAAAGGC 660
Qy 221 SerArgCysTrpGlyLysSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGCTGGGGAGAGATTCTGAGGATTGTGAGAGCTTGCAGCGCTGCTGTGTC 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGGCGCTGCAAGGGGCGCTGCTGAGGATTGTGAGAGCTTGCAGCGCTGCTGTGTC 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTGCACGGGGCCCCAGACCTCTGACTGCTGGCTGGCTGCCCTTCAACACACAGT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCACTCTGTAGCTGCACCTGCCAGCCCTGGTCACTCAACACACAGACACAGTTTGTAGTCC 900
Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320

Db 901 ATGCCCAATCCGAGGCGCGGTATACATTTCGGCCAGCTGTGTGACTGCTGTCCCTAC 960
Qy 321 AsnTyrLeuSerThrAspValGlySerAlaSerIleIleAsnPheGluLysLeuGlyAla 340
Db 961 AACTACCTTTCTACGGACGTGGATCCGCTAGCATCATTAATTTTCGAAAGTGTGGCGCT 1020
Qy 341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyValAspLeu 360
Db 1021 GGGGGCATGTCTCACACAGGACCGCAGCTCATCTACACAGAGTGGCGGTGGGACCTG 1080
Qy 361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
Db 1081 ACCTTAGGGCTGGAGCCCTCTGAAGAGGAGGCCCCAGGTCTCCACTGGCACCCCTCCGAA 1140
Qy 381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
Db 1141 GGGGCTGGCTCCGATGTATTTGATGGTGACCTGGGAATGGGGGCGAGCAAGGGGCTGCAA 1200
Qy 401 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 420
Db 1201 AGCTTCCCAACATGACCCCGCCCTCTACAGGGTACAGTGAAGACCCCCACAGTACCC 1260
Qy 421 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 440
Db 1261 CTGCGCTCTGAGACTGATGGCTACGTTGCCCCCTGACCTGCAGGCCCGCAGCTGAATAT 1320
Qy 441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
Db 1321 GTGAACAGCCAGATGTTTCGGCCCGCAGCCCTTTCGCCCGAGAGGGGCTCTGCTGCT 1380
Qy 461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
Db 1381 GCCCGACCTGCTGGTGCCACTCTGAAAGGGCCAGACTCTCTCCCGAGGAGAGATGGG 1440
Qy 481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 500
Db 1441 GTGCTGAAGACGTTTTTTCCTTTGGGGGTGCCGTGGAGAACCCCGAGTACTTGACACC 1500
Qy 501 GlnGlyGlyAlaAlaProGlnProHisProProProAlaPheSerProAlaPheAspAsn 520
Db 1501 CAGGAGGAGCTGCCCTCAGCCCGCCCTCTCTGCTTTCAGCCCGAGCTTTCAGAAC 1560
Qy 521 LeuTyrTyrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGly 540
Db 1561 CTCTATTAATGCTGGACCCAGACCCAGAGGGGGGGCTCCACCCAGCACCTTCAAGGG 1620
Qy 541 ThrProThrAlaGluAsnProGluTyrLeuGlyLeuAspValProAlaAlaAlaProThr 560
Db 1621 ACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGACGTGCCAGCGCGCCGCCACC 1680
Qy 561 ArgSerProAsnProValThrArgProTrpLysHisValAspAlaIleLysGluAlaLeu 580
Db 1681 CGCTCACCAACCTGTCCCGGGCTTGGAGGATGTAGATGCCATCAAGAGACTCTG 1740
Qy 581 SerLeuLeuAsnAspMetArgAlaLeuGluAsnGluLysAsnGluAspValAspIleIle 600
Db 1741 AGCTTCTTAAATGACATGCGTCTCTGGAGAACGAAAGAACGAAAGACGTAGACATCATC 1800
Qy 601 SerAsnGluPheSerIleGlnArgProThrCysValGlnThrArgLeuLysLeuTyrLys 620
Db 1801 TCTAATGATTTCTCATCCAGGCGGCACATGTGTGCGAGCCCGCTGAAAGCTATACAAG 1860
Qy 621 GlnGlyLeuArgGlyAsnLeuThrLysLeuAsnGlyAlaLeuThrMetIleLeuAsnHis 640
Db 1861 CAGGCTCTACGGGGCAACCTCACAACTCAATGGCGCTTGACCATGATAGCCAGCCAC 1920
Qy 641 TyrGlnThrAsnCysProProThrProGluThrAspCysGluIleGluValThrThrPhe 660
Db 1921 TACCAGAGACTGCGCTCCAAACCCCGAAACTGACTGTGAATAGAGTCAACACCTTT 1980
Qy 661 GluAspPheIleLysAsnLeuLysGlyPheLeuPheAspIleProPheAspCysTrpLys 680

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 686.431 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLAASLSGLFLF.....CWKPVOKGAPPPPAHHHHH 697

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-O=/cpn2.1/USPTO spool/US09821883/runat_09092004.105126.6623/app_query.fasta_1.3100

-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09821883 @CNC 1.1646 @runat_09092004.105126.6623 -NCPU=6 -ICPU=3

-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002s:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3783	100.0	2091	AAD21567	Aad21567 Human HER
2	3473.5	91.8	2070	AAD21565	Aad21565 Human HER
3	3014	79.7	1692	AAD21566	Aad21566 Human HER
4	2957.5	78.2	1665	AAD21564	Aad21564 Human HER
5	2478	65.5	1437	AAD21568	Aad21568 Human HER
6	2439.5	64.5	4530	AAT01585	Aat01585 Her-2/neu
7	2439.5	64.5	4530	AAT71253	Aat71253 Human HER
8	2439.5	64.5	4530	AAZ60815	Aaz60815 Nucleotid

9	2439.5	64.5	4530	4	AAD19731	Aad19731 Human tyz
10	2439.5	64.5	4530	6	ABN85585	Abn85585 Human HER
11	2439.5	64.5	4530	6	ABZ35012	Abz35012 Human gen
12	2439.5	64.5	4530	6	ABV94128	Abv94128 Breast ca
13	2439.5	64.5	4530	6	ABK83918	Abk83918 Human CDN
14	2439.5	64.5	4530	7	ACC50139	Acc50139 Breast ca
15	2439.5	64.5	4530	7	ABQ83856	Abq83856 Human Her
16	2439.5	64.5	4530	8	AAD58073	Aad58073 Human c-e
17	2439.5	64.5	4530	9	ADC09594	Adc09594 Her2/Neu
18	2439.5	64.5	4530	2	ABQ76220	Abq76220 Human tum
19	2439.5	64.4	4473	2	AAZ31071	Aaz31071 HER-2 nuc
20	2439.5	64.4	4473	6	ABZ34969	Abz34969 Human gen
21	2439.5	64.4	4473	6	AAD38904	Aad38904 Human Her
22	2439.5	64.4	4473	7	ACC69999	Acc69999 Human c-e
23	2439.5	64.4	4473	7	ADC35148	Adc35148 Human bre
24	2439.5	64.4	4473	9	ADD25483	Add25483 Binding d
25	2439.5	64.3	9274	6	ABK14057	Abk14057 Human HER
26	2439.5	64.3	9274	6	ABK14057	Abk14057 Human HER
27	2439.5	64.3	4472	3	AAI14812	Aai14812 CDNA enco
28	2439.5	64.2	9274	4	AAF24297	Aaf24297 HER2 tran
29	2439.5	64.0	3765	9	ADB67620	Adb67620 Human epi
30	2439.5	64.0	3768	3	AAA09455	Aaa09455 Human her
31	2439.5	64.0	3768	6	ABV78168	Abv78168 Human ERB
32	2439.5	64.0	3768	6	ABZ35744	Abz35744 Human ERB
33	2439.5	64.0	3768	6	AAD43935	Aad43935 Human HER
34	2439.5	64.0	3768	6	ABX09987	Abx09987 Human ERB
35	2439.5	64.0	3768	6	AAD43986	Aad43986 Human Her
36	2439.5	64.0	3768	6	ABK14058	Abk14058 Human pol
37	2439.5	64.0	3768	6	ABL91709	AbL91709 Human pol
38	2439.5	64.0	3768	7	ACC57649	Acc57649 Human pro
39	2439.5	63.9	3768	2	AAT40739	Aat40739 HER-2/neu
40	2439.5	63.9	3768	2	AAH23392	Aah23392 Human HER
41	2439.5	63.9	3768	5	AAH23392	Aah23392 Human HER
42	2439.5	63.9	3768	6	ABK10730	Abk10730 Human Her
43	2439.5	63.9	3768	6	ABK10730	Abk10730 Human Her
44	2439.5	63.9	3768	6	ABA92350	AbA92350 Human Her
45	2439.5	63.8	4606	9	ADB47370	Adb47370 Human CDN

ALIGNMENTS

RESULT 1

AAD21567

ID AAD21567 standard; DNA; 2091 BP.

XX AC AAD21567;

XX AC AAD21567;

DT 28-JAN-2002 (first entry)

XX DE Human HER500-rGM-CSF fusion DNA construct comprising OVA-derived peptide.

XX DE Immunostimulatory fusion protein; IFP; antigen component; therapy;

XX DE Immunostimulatory component; T-cell mediated immune response; DC;

XX DE Dendritic cell; colon cancer; breast carcinoma; ovarian cancer;

XX DE PAP protein; Ala Arg linker; membrane distal extracellular domain;

XX DE Membrane distal intracellular domain; C-terminal tag; human; GM-CSF;

XX DE HER-2 protein; granulocyte-macrophage colony stimulating factor;

XX DE Ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion DNA; da.

XX OS Homo sapiens.

XX OS Rattus norvegicus.

XX OS Unidentified.

XX OS Chimeric.

XX Key Location/Qualifiers

XX CDS 1..2091

XX FT /*tag= a

XX FT /product= "Human HER500-rGM-CSF fusion protein construct

XX FT comprising human PAP signal sequence, mature PAP protein,

XX FT an Ala Arg linker, human HER-2 signal sequence, mature

XX FT HER-2 membrane distal extracellular domain, an Ala

XX FT linker, an ovalbumin (OVA)-derived immunodominant

XX FT octapeptide, HER-2 membrane distal intracellular domain,

XX FT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 7467.3 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARASLSGLFLF.....CWKPVQKGAPPPPAHHHHH 697

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlh
-Q/cnr2_1/USPTO_spool/US09821883/runat_09092004_105127_6631/app_query.fasta_1.3100
-DB=GenEmbl -QFMT=fastap -SUFFIX=rg -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CNR 1 1 16795 @runat_09092004_105127_6631 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.ev.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sv.*
39: em.hgo.hum.*
40: em.hgo.mus.*
41: em.hgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3783	100.0	2091	6	AX268288	Sequence
2	3473.5	91.8	2070	6	AX268286	Sequence
3	3014	79.7	1692	6	AX268287	Sequence
4	2957.5	78.2	1665	6	AX268285	Sequence
5	2478	65.5	1437	6	AX268289	Sequence
6	2439.5	64.5	4530	6	I21124	Sequence 9
7	2439.5	64.5	4530	6	I59745	Sequence 9
8	2439.5	64.5	4530	6	AR202597	Sequence
9	2439.5	64.5	4530	6	AR283481	Sequence
10	2439.5	64.5	4530	6	AR344811	Sequence
11	2439.5	64.5	4530	6	AX282577	Sequence
12	2439.5	64.5	4530	6	AX587649	Sequence
13	2439.5	64.5	4530	6	AX544071	Sequence
14	2439.5	64.5	4530	6	AX771418	Sequence
15	2439.5	64.5	4530	6	BD005474	Cellular
16	2439.5	64.5	4530	9	HUMHER2A	M11730 Human tyros
17	2434.5	64.4	4473	6	AR080259	Sequence
18	2434.5	64.4	4473	6	AR167390	Sequence
19	2434.5	64.4	4473	6	AR392088	Sequence
20	2434.5	64.4	4473	9	HSERB2R	Sequence
21	2434	64.3	9274	6	AR409602	Sequence
22	2434	64.3	9274	6	AX060703	Sequence
23	2422	64.0	3768	6	AR409603	Sequence
24	2422	64.0	3768	6	AX060704	Sequence
25	2422	64.0	3768	6	AX467229	Sequence
26	2422	64.0	3768	6	AX481438	Sequence
27	2422	64.0	3768	6	BD224136	Novel met
28	2417	63.9	3768	6	AR034479	Sequence
29	2417	63.9	3768	6	BD267514	HER-2/neu
30	2417	63.9	3768	6	AX201817	Sequence
31	2417	63.9	3768	6	AX380923	Sequence
32	2417	63.9	3768	6	AX384604	Sequence
33	2417	63.9	3768	6	AX465456	Sequence
34	2346	62.0	3678	6	AX505114	Sequence
35	2142.5	56.6	2763	6	AX380942	Sequence
36	2142.5	56.6	2763	6	AX380944	Sequence
37	2136	56.5	3780	4	AB008451	Canis fam
38	2033.5	53.8	2385	6	AR082744	Sequence
39	2033.5	53.8	2385	6	AR099963	Sequence
40	2033.5	53.8	2385	6	AR143949	Sequence
41	1983.5	52.4	4506	10	AK129487	Mus muscu
42	1976.5	52.2	4727	10	AK116182	Rattus no
43	1976	52.2	3955	6	BD267515	HER-2/neu
44	1976	52.2	3955	6	I21129	Sequence 14
45	1976	52.2	3955	6	I59750	Sequence 14

ALIGNMENTS

RESULT 1

AX268288 2091 bp DNA linear PAT 29-OCT-2001
LOCUS Sequence 9 from Patent WO0174855.
ACCESSION AX268288
VERSION AX268288.1 GI:16541540
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1
REFERENCE
AUTHORS Laus, R., Vidovic, D. and Graddis, T.
TITLE Compositions and methods for dendritic cell-based immunotherapy
JOURNAL Patent: WO 0174855-A 9 11-OCT-2001;
DENDREON CORPORATION (US)
FEATURES
source
Location/Qualifiers
1..2091
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HER500*-rGM-CSF construct"

ORIGIN
Alignment Scores:
Pred. No.: 1.79e-139 Length: 2091
Score: 3783.00 Matches: 697
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-821-883-4 (1-697) x AX268288 (1-2091)

QY 1 MetArgAlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
DB 1 ATGAGAGTGCACCCCTCTCTGCGCCAGGCGAGCAAGCCTTAGCCTTGGCTTCTGTGTT 60
QY 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
DB 61 CTGCTTTTTCCTGCTAGACCGAGTGTACTAGCCNAGGAGTTGGCGCGGGCCGG 120
QY 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
DB 121 TCGACCCCAAGTGTGCACCGGCACAGACATGAAGCTCGCGCTCCCTGCCAGTCCCGAGACC 180
QY 61 HisLeuAspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlyAsnLeuGlu 80
DB 181 CACCTGGACATGCTCCGCCACCTCTACACAGGCTGCGAGTGGTGAGGAAACCTGGAA 240
QY 81 LeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln 100
DB 241 CTCACCTACCTGCCCAACCAATGCCAGCTGTCTTCTTGCAGATATCCAGGAGGTGCAG 300
QY 101 GlyTyrrValLeuIleAlaHisasnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
DB 301 GGCCTACGTGCTCATCGCTCACCAACAGTGAAGCAGGTCCCACTGCAGAGGCTGCGGATT 360
QY 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGlyAsp 140
DB 361 GTGCCAGGACCACCTCTTTGAGGACAACTATATGCTGCGCTGCTAGACAAATGGAGAC 420
QY 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
DB 421 CGCTGGAACATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTGGGGAGGTGCAG 480
QY 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
DB 481 CTTCAAGCCTCACAGATCTTGAAGAGGGGTCTTGATCCAGCGGAAACCCCAAGCTC 540
QY 181 CysTyrrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
DB 541 TGCTACCAAGGACCATGATTGTGGAGGACATCTTCCCAAGAACCAACAGCTGCTCTC 600
QY 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220

601 ACACGTAGACACCAACCGCTCTCGGCTCGCCACCCCTGTTCTCCGATGTGTAGGGC 660
221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
661 TCCCGCTGCTGGGAGAGAGTTCTCAGGATTGTTCAGAGCCTCAGCGCATCTCTGTGTC 720
241 GlyGlyCysAlaAatqCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
721 GGTGCTGTGCCCTGCAAGGGGCCACTGCCCATCTGACTGCTGCATGAGGAGGTGTGT 780
261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
781 GCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGluSer 300
841 GGCATCTGTGAGCTGCATGCGCCAGCTGCGCCAGCTGCGCCAGCTGCGCCAGCTGCGCC 900
301 MetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysProTyrr 320
901 ATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCTGCT 960
321 AsnTyrrLeuSerThrAspValGlySerAlaSerIleIleAsnPheGluLysLeuGlyAla 340
961 AACTACCTTTCTACGAGCTGGGATCCGCTAGCATCATTAATTCGAGAGTGTGGCGCT 1020
341 GlyGlyMetValHisArgHisArgSerSerSerThrArgSerGlyGlyCysAspLeu 360
1021 GGGGCATGGTCCACCAAGGCGCGCTATACATTCGGCGCCAGCTGTGTGACTGCTGCTGCT 1080
361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
1081 ACATAGGGCTGGAGCCTCTGAAGAGGAGGCGCCAGCTTCCACTGGCACCTCCGNA 1140
381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
1141 GGGGCTGCTCCGATGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
401 SerLeuProThrHisAspProSerProLeuGlnArgTyrrSerGluAspProThrValPro 420
1201 AGCTCCCAACACATGACCCAGCCTCTACAGCGGTACAGTGAAGAGCCCAAGTACCC 1260
421 LeuProSerGluThrAspGlyTyrrValAlaProLeuThrCysSerProGlnProGluTyrr 440
1261 CTGCTCTGAGACTGATGGCTACGTTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
1321 GTGAACACGACAGATGTTCCGCCCCAGCCTTCCGCCCGAGAGGGGCTCTGCTGCTGCT 1380
461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
1381 GCCGACCTGTGTGGTCCACTGTGAAGAGGCGCAAGACTCTCTCCCAAGGGAAGATGGG 1440
481 ValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrrLeuThrPro 500
1441 GTGCTCAAGACGTTTTTGGCTTTGGGGGTGCGTGGAGAACCCCGAGTACTTGCACCC 1500
501 GlnGlyGlyAlaAlaProGlnProHisProProAlaPheSerProAlaPheAspAsn 520
1501 CAGGAGGAGCTGCCCTCAGCCCCCAGCCTCTCTGCTGCTTCCAGCCAGCCTTCGACAC 1560
521 LeuTyrrTyrrTrpAspGlnAspProGluArgGlyAlaProProSerThrPheLysGly 540
1561 CTCTATTACTGGGACCGAGCCACAGAGGGGGGCTCCACCCAGCAGCTTCAAGAGG 1620
541 ThrProThrAlaGluAsnProGluTyrrLeuGlyLeuAspValProAlaAlaAlaProThr 560
1621 ACACCTACGACAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGCGGCGCCGCCACC 1680
561 ArgSerProAsnProValThrArgProTrpLysHisValAspAlaIleLysGluAlaLeu 580

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 69.255 Seconds
(without alignment)
3175.456 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLAARASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rhodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	56.5	1259	6 O18735	O18735 canis fam1
2	1962.5	51.9	1259	11 Q8K3P9	Q8K3P9 rattus norv
3	1612.5	42.6	419	4 Q8UKY9	Q8UKY9 homo sapien
4	1380	36.5	711	11 Q80Y89	Q80Y89 mus musculu
5	1014	26.8	881	11 Q8C057	Q8C057 mus musculu
6	966	25.5	367	11 Q8R2X1	Q8R2X1 mus musculu
7	906	23.9	165	4 Q14256	Q14256 homo sapien
8	894.5	23.6	412	4 Q8WYV0	Q8WYV0 homo sapien
9	859	22.7	431	13 Q7SY19	Q7SY19 brachydanio
10	751	19.9	527	13 Q90836	Q90836 gallus gall
11	723	19.1	149	6 Q8EG66	Q8EG66 oryctolagus
12	717.5	19.0	643	11 Q9ERV6	Q9ERV6 mus musculu
13	717.5	19.0	655	11 Q9WVF5	Q9WVF5 mus musculu
14	717.5	19.0	1210	11 Q9EP98	Q9EP98 mus musculu
15	713.5	18.9	1209	11 Q9QX70	Q9QX70 rattus norv
16	708.5	18.7	1209	6 Q8MIL8	Q8MIL8 sus scrofa

ALIGNMENTS

RESULT 1

ID O18735 PRELIMINARY; PRT; 1259 AA.
AC O18735;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE ErBB-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_taxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "cDNA cloning of erbB-2 from canine mammary gland.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1; -.
DR HSSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_Fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP_motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF000069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.

Q7szf7 brachydanio
Q9ese0 rattus norv
Q9bud7 homo sapien
Q9yh40 xiphophorus
Q8aw81 brachydanio
Q9ph22 gallus gall
P79754 fugu rubrip
Q8vh40 sigmodon hi
Q8mlw0 drosophila
Q8bi9h anopheles g
Q8enz2 drosophila
Q8cfb5 marionnes un
Q8myk4 ovis aries
Q8wn17 equus caball
Q95l10 equus caball
Q9gl44 macaca mula
Q7yrf7 felis silve
Q85y5 papio anubi
Q9w6f6 gallus gall
Q9bg64 oryctolagus
Q9bg65 oryctolagus
Q23821 caenorhabdi
Q26569 schistosoma
Q26566 schistosoma
Q26567 schistosoma
Q26568 schistosoma
Q9y1x8 ephydatia f
Q99j91 marmota mon
Q86md7 echinococcu

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 12.7941 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLFLF.....CWKPVQKGAPPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2417	63.9	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1959	51.8	1257	1 ERB2_RAT	P06494 rattus norv
3	1946.5	51.5	1254	1 ERB2_MESAU	Q00553 mesocricetu
4	752	19.9	703	1 EGFR_CHICK	P13387 gallus gall
5	723.5	19.1	1308	1 ERB4_HUMAN	Q15303 homo sapien
6	719.5	19.0	1210	1 EGFR_HUMAN	P00533 homo sapien
7	719.5	19.0	1308	1 ERB4_RAT	Q02956 rattus norv
8	717.5	19.0	1210	1 EGFR_MOUSE	Q01279 mus musculu
9	697	18.4	1342	1 ERB3_HUMAN	P21860 homo sapien
10	683	18.1	127	1 CSF2_RAT	P48750 rattus norv
11	668	17.7	1339	1 ERB3_XIPMA	Q62799 rattus norv
12	651	17.2	1167	1 XMRK_XIPMA	P13388 xiphophorus
13	539.5	14.3	1426	1 EGFR_DROME	P04412 drosophila
14	466	12.3	141	1 CSF2_MOUSE	P01587 mus musculu
15	461	12.2	144	1 CSF2_SHEEP	P28773 ovis aries
16	444	11.7	144	1 CSF2_CEREL	P51748 cervus elap
17	442	11.7	144	1 CSF2_HUMAN	P04141 homo sapien
18	421.5	11.1	143	1 CSF2_BOVIN	P11052 bos taurus
19	417	11.0	144	1 CSF2_PIG	Q29118 sus scrofa
20	404.5	10.7	1367	1 Lf23_CABEL	P24348 caenorhabdi
21	396.5	10.5	144	1 CSF2_CANFA	P48749 canis famil
22	388	10.3	144	1 CSF2_FELCA	O62757 felis silve
23	366.5	9.7	140	1 CSF2_CAVPO	Q00481 cavia porce
24	318	8.4	1363	1 ILPR_BRALA	O02466 branchiosto
25	301.5	8.0	245	1 ERB2_MOUSE	P70424 mus musculu
26	276.5	7.3	1477	1 HYTK_HYDAT	Q05197 hydra atten
27	271	7.2	2146	1 INSR_DROME	P09208 drosophila
28	258	6.8	1382	1 INSR_HUMAN	P06213 homo sapien
29	252.5	6.7	1372	1 INSR_MOUSE	P15208 mus musculu
30	252	6.7	1607	1 MIPR_LYMST	Q25410 lynnaea sta
31	251.5	6.6	1300	1 IRR_MOUSE	Q9wt14 mus musculu
32	249	6.6	1383	1 INSR_RAT	P15127 rattus norv
33	248.5	6.6	1297	1 IRR_HUMAN	P14616 homo sapien

RESULT 1

ID	ERB2_HUMAN	STANDARD;	PRT;	1255 AA.
AC	P04626;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).			
DE	ERBB2 OR HER2 OR NGL OR NEU.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RA	Saito T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erb-B-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
RN	(2)			
RP	SEQUENCE FROM N.A., AND VARIANT ALA-1170.			
RX	MEDLINE=86070181; PubMed=2999974;			
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,			
RA	Franccke U., Levinson A., Ullrich A.;			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
RN	(3)			
RP	SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.			
RX	Rieder M.J., Livingston R.J., Daniele M.R., Montoya M.A., Chung M.-W.,			
RA	Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,			
RA	Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.			
RN	(4)			
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=2995967;			
RA	Semba K., Kanata N., Toyoshima K., Yamamoto T.;			
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RL	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
RN	(5)			
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-1- FUNCTION: Essential component of a neuregulin-receptor complex,			
CC	although neuregulins do not interact with it alone. GP30 is a			
CC	potential ligand for this receptor. Not activated by EGF, TGF-			

34	246.5	6.5	1300	1	IRR_CAVPO	P14617 cavia porce
35	238.5	6.3	1367	1	IG1R_HUMAN	P08069 homo sapien
36	236.5	6.3	1370	1	IG1R_RAT	P24062 rattus norv
37	234.5	6.2	1373	1	IG1R_MOUSE	Q60751 mus musculu
38	228.5	6.0	581	1	IRR_FAT	Q64716 rattus norv
39	207	5.5	1390	1	INSR_AEDAE	Q93105 aedes aegyp
40	168	4.4	386	1	PPAP_HUMAN	P15309 homo sapien
41	165	4.4	5262	1	MLL2_HUMAN	O14686 homo sapien
42	164	4.3	707	1	SPQ_HUMAN	P23246 homo sapien
43	155.5	4.1	497	1	WAS2_HUMAN	Q9v6w5 homo sapien
44	154.5	4.1	1321	1	IRS2_MOUSE	P81122 mus musculu
45	150	4.0	2715	1	MLL4_HUMAN	Q9umh6 homo sapien

ALIGNMENTS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 22.2506 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLF.....CWKPVQKAGPPPAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2417	63.9	1255	1 A24571	protein-tyrosine k
2	1962	51.9	1260	1 TVRTNU	protein-tyrosine k
3	1946.5	51.5	1254	2 I48161	p-185 precursor -
4	752	19.9	1223	1 TVCHLV	epidermal growth f
5	751	19.9	527	2 A42032	epidermal growth f
6	723.5	19.1	1308	2 A47253	epidermal growth f
7	719.5	19.0	1210	1 GQHUE	epidermal growth f
8	717.5	19.0	1210	2 A53183	epidermal growth f
9	713.5	18.9	644	2 A36325	epidermal growth f
10	697	18.4	1342	2 A36223	kinase-related tra
11	683	18.1	127	2 I46289	granulocyte-macrop
12	662	17.5	1339	2 J4387	epidermal growth f
13	651	17.2	1166	1 S06142	protein-tyrosine k
14	539.5	14.3	843	2 A27131	epidermal growth f
15	473	12.5	153	1 F0MSGM	granulocyte-macrop
16	461	12.2	144	2 JH0469	granulocyte-macrop
17	457	12.1	144	1 A61632	granulocyte-macrop
18	442	11.7	144	1 F0HUGM	granulocyte-macrop
19	421.5	11.1	143	1 F0B0GM	granulocyte-macrop
20	404.5	10.7	1323	2 E88257	protein let-23 (im
21	404.5	10.7	1374	2 S70712	protein-tyrosine k
22	396.5	10.5	144	2 A44936	granulocyte-macrop
23	390.5	10.3	1369	2 S70713	protein-tyrosine k
24	379	10.0	1330	1 G0FFE	epidermal growth f
25	341	9.0	366	2 D45558	epidermal growth f
26	341	9.0	1717	1 A45558	epidermal growth f
27	331	8.7	333	2 B45558	epidermal growth f
28	331	8.7	342	2 C45558	epidermal growth f
29	318	8.4	1363	2 T43220	insulin-like growth

RESULT 1

A24571

protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human

N:Alternate names: C-erb-B-2 protein precursor; kinase-related transforming protein erbB

C:Species: Homo sapiens (man)

C>Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999

C:Accession: A24571; A25491; A44188; B44188; I59509; I57622

R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;

Nature 319, 230-234, 1986

A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f

A:Reference number: A24571; MUID:86118663; PMID:3003577

A:Accession: A24571

A:Molecule type: mRNA

A:Residues: 1-1255 <YAM>

A:Cross-references: GB:X03363; NID:G31197; PIDN:CAA27060.1; PID:G31198

R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.

Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985

A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm

A:Reference number: A25491; MUID:86016729; PMID:2995967

A:Accession: A25491

A:Molecule type: DNA

A:Residues: 737-1031 <SEM>

A:Cross-references: GB:M11767; NID:G182163; PIDN:AAA35808.1; PID:G553282

R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.

Science 230, 1132-1139, 1985

A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos

A:Reference number: A44188; MUID:86070181; PMID:2999974

A:Accession: A44188

A:Molecule type: DNA

A:Residues: 740-910 <COU1>

A:Cross-references: GB:M12036; NID:G183988; PIDN:AAA35978.1; PID:G183989

A:Accession: B44188

A:Molecule type: mRNA

A:Residues: 1-517; RALL, 522, 'S', 524-654; 'V', 656-1169; 'A', 1171-1255 <COU2>

A:Cross-references: GB:M11730; NID:G183986

R:King, C.R.; Kraus, M.H.; Aaronson, S.A.

Science 229, 974-976, 1985

A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.

A:Reference number: I59509; MUID:85272597; PMID:2992089

A:Accession: I59509

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 832-909 <REX>

A:Cross-references: GB:I29395; NID:G459807; PIDN:AAA35809.1; PID:G459808

R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.

Mol. Cell. Biol. 7, 2597-2601, 1987

A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional

A:Reference number: I57622; MUID:87286899; PMID:3039351

A:Accession: I57622

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-191 <TAL>

A;Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C;Genetics:
A;Gene: GDB:ERBB2; NGL; NEU; HER-2
A;Cross-references: GDB:120613; OMIM:164870
A;Map position: 17q21.1-17q21.1
A;Introns: 25/1; 75/3; 147/1; 883/3
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F;22-653/Domain: extracellular #status predicted <EXT>
F;70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F;395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F;654-675/Domain: transmembrane #status predicted <TM>
F;676-1255/Domain: intracellular #status predicted <INT>
F;718-983/Domain: protein kinase homology <KIN>
F;726-734/Region: protein kinase ATP-binding motif
F;68,124,187,259,530,571,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;753/Active site: Lys #status predicted
F;1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 63.9%; Score 2417; DB 1; Length 1255;
Best Local Similarity 41.2%; Pred. No. 5.6e-129;
Matches 511; Conservative 4; Mismatches 6; Indels 718; Gaps 2;

Qy	35	LARGAASVCTGTDKMLRSPASPTHLDMLRHLRYQGVVQGNLELTYPNTASLSFLQ	94
Db	16	LPPGAASVCTGTDKMLRSPASPTHLDMLRHLRYQGVVQGNLELTYPNTASLSFLQ	75
Qy	95	DIQVQGVVLAHNOVROVPLQRLIRVGTQLFEDNALVALDNGDPLNNTPTVTGASPG	154
Db	76	DIQVQGVVLAHNOVROVPLQRLIRVGTQLFEDNALVALDNGDPLNNTPTVTGASPG	135
Qy	155	GLRELQRLSTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALTLIDNRSRACHPC	214
Db	136	GLRELQRLSTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALTLIDNRSRACHPC	195
Qy	215	SPMCKGSRGWGESSEDCOSLTRVCAGCARGCKPLPTDCHEOCACGCTGPKHSDCLAC	274
Db	196	SPMCKGSRGWGESSEDCOSLTRVCAGCARGCKPLPTDCHEOCACGCTGPKHSDCLAC	255
Qy	275	LHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGSASII	333
Db	256	LHFNHSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGSCTLVC	315
Qy	334	-----	333
Db	316	PLHNQVETABDGTQCEKCKPCARVCYGLGMEHLREVRVTSANIQEFAGCKKIFGSLA	375
Qy	334	-----	333
Db	376	FLPESFDGPASNTAPLOEQVLFETLEITGYLYISAWPDSLPLDSVFONLQVIRGRI	435
Qy	334	-----	333
Db	436	LHNGAYSILTLOGLIGLSWLSRLSRLSGLALIHNTLHCFVHTVPWDLFRNPHQALLH	495
Qy	334	-----	333
Db	496	TANRPEDECVGEGLAGLACHQARGHGWGPGTQCVNCSQFLRGQECVBEVCRLQGLPREYV	555
Qy	334	-----	333
Db	556	NARHCLPCHPCQPONGSVTCFGEADQCVAACHYKDDPPFCVAKCPGKVPDLSYMPIWK	615
Qy	334	-----	333

Db	616	FPDEGACQPCPINCTHSCVDLDDKGCAPABORASPLTIIISAVVGIILLVVLGVVFGILI	675
Qy	334	-----	333
Db	676	KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRIILKETELRKVKVLGSGAFGTIV	735
Qy	334	-----	333
Db	736	KGWIWPDGENVKIPVAIKVIRENTSPKANKEILDEAYMAGVGSPPYVSRLLGICLTSTVQ	795
Qy	334	-----	333
Db	796	LVTQLMEYGLLDHVRENRLGSDLLNWCMIKAGMSYLEVRLVHRDLAARNVLVKS	855
Qy	334	-----	333
Db	856	PNHVKITDFGLARLLDIDETEHADGKVPKMWMALESILRRRTHOSDVWSYGVTVWEL	915
Qy	334	-----	338
Db	916	MTFGAKPYDGIPIAREIPDLLEKGERLPPICTIDVTMVMVKMWIDSECRPRPRELVSE	975
Qy	339	-----	338
Db	976	FSRMARDPQRVVIQNEIDLGPASPLDSTFYRSLLDDDDMGDLVDABEYLVPOQGFCDP	1035
Qy	339	-----	396
Db	1036	APGAGMVHRRSSSTRSGGDLTLGLEPSESEAPRSLAPSGAGSDVFDGDLGMGAA	1095
Qy	397	KGLOSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQVPRPPSPREG	456
Db	1096	KGLOSLPHTDPSPLQRYSEDPTVPLPSETDGYVAPLTCSPQPEYVNPQVPRPPSPREG	1155
Qy	457	PLPAARPAATLRAKTLSPGKGVKVDVAFGAVENPEYLTPOGGAAPQHPHPPAFSP	516
Db	1156	PLPAARPAATLRAKTLSPGKGVKVDVAFGAVENPEYLTPOGGAAPQHPHPPAFSP	1215
Qy	517	AFNLVYWDODPPERGAPPSTFKGTPTAENPEYLGLDVP	555
Db	1216	AFNLVYWDODPPERGAPPSTFKGTPTAENPEYLGLDVP	1254

RESULT 2
TVRTNU
Protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C;Accession: A24562; A61204
R;Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A;Reference number: A24562; UID:86118662; PMID:3945311
A;Molecule type: mRNA
A;Residues: 1-1260 <BAR>
A;Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
R;Masul, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
Carciogenes 12, 1975-1978, 1991
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m
2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A;Reference number: A61204; UID:92035293; PMID:1692063
A;Accession: A61204
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 637-663, 'V', 665-702 <MAS>
A;Note: authors translated the codon GCA for residue 25 as Val
C;Genetics:
A;Gene: neu
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 73.7051 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARAAASLSGLFLF.....CWKPVQKGA PPPRAHHHHH 697

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3783	100.0	697	9	US-09-821-883-4
2	3473.5	91.8	690	9	US-09-821-883-2
3	3014	79.7	564	9	US-09-821-883-3
4	2957.5	78.2	555	9	US-09-821-883-1
5	2585	68.3	919	9	US-09-854-356-6
6	2478	65.5	479	9	US-09-821-883-5
7	2422	64.0	1255	9	US-09-811-123-9
8	2422	64.0	1255	9	US-09-811-115-3
9	2422	64.0	1255	9	US-09-769-508-2
10	2422	64.0	1255	12	US-09-984-092-4
11	2422	64.0	1255	14	US-10-177-293-126
12	2422	64.0	1255	14	US-10-207-498-6
13	2422	64.0	1255	14	US-10-338-730-2
14	2422	64.0	1255	14	US-10-322-892-4
15	2422	64.0	1255	15	US-10-272-437A-28

16	2422	64.0	1255	15	US-10-117-937-594	Sequence 594, App
17	2422	64.0	1255	15	US-10-435-696-36	Sequence 36, Appl
18	2422	64.0	1255	16	US-10-441-779C-4	Sequence 4, Appl
19	2422	64.0	1255	16	US-10-734-564-126	Sequence 126, App
20	2417	63.9	1255	9	US-09-854-356-1	Sequence 1, Appl
21	2417	63.9	1255	9	US-09-930-125-2	Sequence 2, Appl
22	2417	63.9	1255	10	US-09-441-411-6	Sequence 6, Appl
23	2417	63.9	1255	12	US-10-469-162-3	Sequence 3, Appl
24	2417	63.9	1255	12	US-10-253-286-553	Sequence 553, App
25	2417	63.9	1255	12	US-09-765-973-2	Sequence 2, Appl
26	2417	63.9	1255	12	US-10-418-027-3	Sequence 3, Appl
27	2417	63.9	1255	14	US-10-207-555-45	Sequence 45, Appl
28	2417	63.9	1255	14	US-10-313-644-2	Sequence 2, Appl
29	2417	63.9	1255	15	US-10-394-322A-17	Sequence 17, Appl
30	2417	63.9	1255	15	US-10-245-871-553	Sequence 553, App
31	2417	63.9	1255	15	US-10-149-138-4641	Sequence 4641, Ap
32	2417	63.9	1255	16	US-10-647-005-68	Sequence 68, Appl
33	2417	63.9	1255	16	US-10-149-138-4641	Sequence 4641, Ap
34	2410	63.7	1253	14	US-10-146-473-72	Sequence 72, Appl
35	1974.5	52.2	1256	9	US-09-854-356-14	Sequence 118, App
36	1962	51.9	1260	9	US-09-870-759-118	Sequence 118, App
37	1962	51.9	1260	10	US-09-751-708A-118	Sequence 2, Appl
38	1959.5	51.8	1255	9	US-09-854-356-2	Sequence 1, Appl
39	1614.5	42.7	645	9	US-09-921-161-1	Sequence 13, Appl
40	1614.5	42.7	645	14	US-10-268-501-13	Sequence 13, Appl
41	1614.5	42.7	645	15	US-10-608-626-13	Sequence 3, Appl
42	1614.5	42.7	653	9	US-09-854-356-3	Sequence 3, Appl
43	1614.5	42.7	685	15	US-10-412-804A-4	Sequence 4, Appl
44	1614.5	42.7	690	15	US-10-412-804A-11	Sequence 11, Appl
45	1614.5	42.7	712	9	US-09-854-356-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

Query Match 100.0%; Score 3783; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 2e-258;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRAAPLLARAAASLSGLFLFPWLDRSVLAKELARGAASSTQVCTGDMKRLPASSET	60
Db	1	MRAAPLLARAAASLSGLFLFPWLDRSVLAKELARGAASSTQVCTGDMKRLPASSET	60
Qy	61	HLDMRLHYQCQVQVQGNLELTPTNASLSFLQDIEQVGVLIHNVQVPLQRLRI	120
Db	61	HLDMRLHYQCQVQVQGNLELTPTNASLSFLQDIEQVGVLIHNVQVPLQRLRI	120
Qy	121	VRGTQLPEDNYALAVLDNGDPLNNTTPVTGASPGGLRELQRLSUTETILKGGVLIQRNPOL	180

```
Db 121 VRGTLQFEDYALAVLDNGDPLNNTTPTVGTASPGGLRELQLRSITEILKGGVLQRPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALFLIDNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALFLIDNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHECAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGPLPTDCCHECAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Qy 301 MPNPEGRYTFGASCVTACPNYNYLSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Db 301 MPNPEGRYTFGASCVTACPNYNYLSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Qy 361 TLGLEPSEEEAPRPLAPSEAGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPTVP 420
Db 361 TLGLEPSEEEAPRPLAPSEAGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPTVP 420
Qy 421 LPSETDGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Qy 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLYYWDQDPPPERGAPPSTFKG 540
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLYYWDQDPPPERGAPPSTFKG 540
Qy 541 TPTAENPEYLGDPVAPAAAPTRSPNPVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Db 541 TPTAENPEYLGDPVAPAAAPTRSPNPVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Qy 601 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTPETDCIEVITTF 660
Db 601 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTPETDCIEVITTF 660
Qy 661 EDFIKNLKGLFDIIPFCWKVPQKGAAPPPAHHHHHH 697
Db 661 EDFIKNLKGLFDIIPFCWKVPQKGAAPPPAHHHHHH 697

RESULT 2
US-09-821-883-2
; Sequence 2, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-HCM-CSF construct
US-09-821-883-2

Query Match 91.8%; Score 3473.5; DB 9; Length 690;
Best Local Similarity 91.7%; Pred. No. 1.4e-236;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;
```

```
Qy 1 MRAAPLLIARAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGTDKMLRLPASPET 60
Db 1 MRAAPLLIARAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGTDKMLRLPASPET 60
```

```
Qy 61 HLDMLRHLYGQCVVQGNLELTYPNTNASISFLQDIOEVQGYVLIHAHQVQFLORLRI 120
Db 61 HLDMLRHLYGQCVVQGNLELTYPNTNASISFLQDIOEVQGYVLIHAHQVQFLORLRI 120
Qy 121 VRGTLQFEDYALAVLDNGDPLNNTTPTVGTASPGGLRELQLRSITEILKGGVLQRPQL 180
Db 121 VRGTLQFEDYALAVLDNGDPLNNTTPTVGTASPGGLRELQLRSITEILKGGVLQRPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALFLIDNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALFLIDNRSRACHPCSPMCKGSRGWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHECAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Db 241 GGCARCKGPLPTDCCHECAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTFES 300
Qy 301 MPNPEGRYTFGASCVTACPNYNYLSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Db 301 MPNPEGRYTFGASCVTACPNYNYLSTDVGSASIIINFEKLGAGGMVHRRHRSSTSRGGDL 360
Qy 361 TLGLEPSEEEAPRPLAPSEAGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPTVP 420
Db 361 TLGLEPSEEEAPRPLAPSEAGSDVFDGDLGMAAGLQSLPHTDPSPLQRYSEDPTVP 420
Qy 421 LPSETDGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVAPLTCSPQPEYVNOQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Qy 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLYYWDQDPPPERGAPPSTFKG 540
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLYYWDQDPPPERGAPPSTFKG 540
Qy 541 TPTAENPEYLGDPVAPAAAPTRSPNPVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Db 541 TPTAENPEYLGDPVAPAAAPTRSPNPVTRPKHVDVDAIKEALSLLNDMRALENEKNEVDII 600
Qy 591 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTPETDCIEVITTF 660
Db 591 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTPETDCIEVITTF 660
Qy 661 EDFIKNLKGLFDIIPFCWKVPQKGAAPPPAHHHHHH 697
Db 661 EDFIKNLKGLFDIIPFCWKVPQKGAAPPPAHHHHHH 697

RESULT 3
US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 79.7%; Score 3014; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 3e-204;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 20.3037 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-4
Perfect score: 3783
Sequence: 1 MRAAPLLARASLSGLFL.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	64.0	1255	4	US-09-527-487-2
2	2422	64.0	1255	4	US-09-811-115-3
3	2417	63.9	1255	1	US-08-467-083-68
4	2417	63.9	1255	1	US-08-414-417B-68
5	2417	63.9	1255	2	US-08-486-348A-68
6	2417	63.9	1255	2	US-08-625-101-2
7	2417	63.9	1255	2	US-08-468-545B-68
8	2417	63.9	1255	2	US-08-356-786-2
9	2417	63.9	1255	3	US-08-466-680B-68
10	2417	63.9	1255	4	US-09-354-533-68
11	2397	63.4	1255	2	US-08-484-438-8
12	2001.5	52.9	782	2	US-09-146-283-4
13	2001.5	52.9	782	3	US-08-579-823A-4
14	2001.5	52.9	782	3	US-09-344-195-4
15	1610.5	42.6	419	4	US-09-630-155-2
16	1599.5	42.3	624	3	US-08-422-108-1
17	1599.5	42.3	624	4	US-08-422-734-1
18	1177	31.1	580	1	US-08-414-417B-69
19	1177	31.1	580	2	US-08-486-348A-69
20	1177	31.1	580	2	US-08-468-545B-69
21	1177	31.1	580	3	US-08-466-680B-69
22	1177	31.1	580	4	US-09-354-533-69
23	860	22.7	166	4	US-09-648-067A-1
24	723.5	19.1	911	2	US-08-484-438-10
25	723.5	19.1	1058	2	US-08-484-438-4
26	723.5	19.1	1308	2	US-08-484-438-2
27	719.5	19.0	644	1	US-08-336-708A-9

28	719.5	19.0	1210	2	US-08-484-438-7	Sequence 7, Appli
29	719.5	19.0	1210	2	US-08-475-035-4	Sequence 4, Appli
30	701.5	18.5	478	4	US-09-570-454-2	Sequence 2, Appli
31	701.5	18.5	478	4	US-09-867-521-2	Sequence 2, Appli
32	697	18.4	1342	1	US-07-978-895-4	Sequence 4, Appli
33	697	18.4	1342	2	US-08-484-438-9	Sequence 9, Appli
34	697	18.4	1342	2	US-08-473-119-4	Sequence 4, Appli
35	697	18.4	1342	2	US-08-475-352-4	Sequence 4, Appli
36	697	18.4	1342	4	US-09-170-699-4	Sequence 4, Appli
37	697	18.4	1343	6	5183884-4	Patent No. 5183884
38	529.5	14.0	515	2	US-09-146-283-2	Sequence 2, Appli
39	529.5	14.0	515	3	US-08-579-823A-2	Sequence 2, Appli
40	529.5	14.0	515	3	US-09-344-195-2	Sequence 2, Appli
41	493	13.0	97	1	US-08-421-356-3	Sequence 3, Appli
42	493	13.0	97	1	US-09-046-783-3	Sequence 3, Appli
43	472.5	12.5	219	2	US-08-902-516-2	Sequence 2, Appli
44	472.5	12.5	219	4	US-09-847-185-2	Sequence 2, Appli
45	469	12.4	141	1	US-08-259-696B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527.487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match	64.0%	Score	2422	DB	4	Length	1255
Best Local Similarity	41.3%	Pred. No.	2.2e-187				
Matches	512	Conservative	4	Mismatches	5	Indels	718
						Gaps	2
Qy	35	LARGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQ	94				
Db	16	LPPGAASQVCTGDMKLRLPASPETHLDMLRHLYQGCVVQGNLELTYLPTNASLSFLQ	75				
Qy	95	DIQEVQGYVLIAHNVQRQVPLQRLRIVRGTLQDPEDNYALAVLDNGDPLNNTTPVTGASPG	154				
Db	76	DIQEVQGYVLIAHNVQRQVPLQRLRIVRGTLQDPEDNYALAVLDNGDPLNNTTPVTGASPG	135				
Qy	155	GLRELQRLSLTEILKGGVLQORNQOLCVQDTILWKDIFHKQNQLALTIDNRSRACHPC	214				
Db	136	GLRELQRLSLTEILKGGVLQORNQOLCVQDTILWKDIFHKQNQLALTIDNRSRACHPC	195				
Qy	215	SPMKGRCRCWGESSEDCQSLTRTVTCAGCARCKGLPTDCCHEQCAAGCTGPKHSDCLAC	274				
Db	196	SPMKGRCRCWGESSEDCQSLTRTVTCAGCARCKGLPTDCCHEQCAAGCTGPKHSDCLAC	255				
Qy	275	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGSASII	333				
Db	256	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGSASII	315				
Qy	334	-----	333				
Db	316	PLHNQVETADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIOFAGCKKIFGSLA	375				
Qy	334	-----	333				
Db	376	FLPESFGDPPASNTAPLQPEQLQVFETLEETGYLISNWPDSLPDLSVFNQLQVIRGRI	435				

```
Qy 334 ----- 333
Db 436 LHNAYSLTLOGLISWLGSLRLSRLGSLALIHNNTHLCFVHTVPWDLFRNPQALLH 495
Qy 334 ----- 333
Db 496 TANRPEDECYVGLACHQLCARGHCWGPGPTQCVCNCSQFLRGQCEBECRVLOGLPREYV 555
Qy 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
Qy 334 ----- 333
Db 616 FPDEGACQPCPNCTHSCVDLDDKGPABORASPLTSIVSADVGLLVVVLGVVFGILI 675
Qy 334 ----- 333
Db 676 KRRQKIRKYTMRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVILGSAFCTVY 735
Qy 334 ----- 333
Db 736 KGIWPDGENVKIPVAIKVLRNTSPKANKEILDEAYVMAGVGSYVSRLLIGICTSTVQ 795
Qy 334 ----- 333
Db 796 LVTQMPYGCLLDHRNENRGLSGDILLNMCQIAKMSYLEVDVLRVHRDLAARNVLKS 855
Qy 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIMWMALESILRRRFTHQSDVMSYGVTVWEL 915
Qy 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVWKMWIDSECRPRELVS 975
Qy 339 ----- 338
Db 976 FSRWAPQRFVVIQNEDELGPASPLDSTFYRSLLDDMDGLVDAEYLVPOQFFCFDP 1035
Qy 339 --GAGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSPAPSEAGSDVPDGLMGAA 396
Db 1036 APGAGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSPAPSEAGSDVPDGLMGAA 1095
Qy 397 KGLQSLTHDPSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVRRPQPSREG 456
Db 1096 KGLQSLTHDPSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVRRPQPSREG 1155
Qy 457 PLPAARAGATLERAKTILSPKNGVWVDVAFGAVENPEYLTPOQGAAPOPHPPPAFSP 516
Db 1156 PLPAARAGATLERAKTILSPKNGVWVDVAFGAVENPEYLTPOQGAAPOPHPPPAFSP 1215
Qy 517 AFDNLYWQDQPPRGPAPPSTFKTPTAENPEYLGLDVP 555
Db 1216 AFDNLYWQDQPPRGPAPPSTFKTPTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match 64.0%; Score 2422; DB 4; Length 1255;
Best Local Similarity 41.3%; Pred. No. 2.2e-187;
Matches 512; Conservative 4; Mismatches 5; Indels 718; Gaps 2;

Qy 35 LARGAASQVCTGTGDMKRLPASPETHLDMRLHYQCGVQVQGNLELYLPTNASLSFLQ 94
Db 16 LPFGAASQVCTGTGDMKRLPASPETHLDMRLHYQCGVQVQGNLELYLPTNASLSFLQ 75
Qy 95 DIQEVQGYVLIANQVROVPLQRLRIVRGTQLEPDNYALAVLDNGDPLNNTTPTVTGASPG 154
Db 76 DIQEVQGYVLIANQVROVPLQRLRIVRGTQLEPDNYALAVLDNGDPLNNTTPTVTGASPG 135
Qy 155 GLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQNLALTLIDTNRSRACHPC 214
Db 136 GLRELQRLSLTEILKGGVLIQRNPQLCYQDTILWKDIFHKNQNLALTLIDTNRSRACHPC 195
Qy 215 SPMCKSRGCGESSEDCQSLTRTVACGGCARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKSRGCGESSEDCQSLTRTVACGGCARCKGPLPTDCHEQCAAGCTGPKHSDCLAC 255
Qy 275 LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGCACVTPYNYLSTDVGSASII- 333
Db 256 LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGCACVTPYNYLSTDVGSCTLVC 315
Qy 334 ----- 333
Db 316 PLHNQEVTAEDGTQRCCKSPCARVCYGLGMEHLREVRAVTSANIOEPAGCKKIFGSLA 375
Qy 334 ----- 333
Db 376 FLPEFDGDPASNTAPLQPEQLQVFETLEBITGYLISAWPDSLPDLSVFNQLQVIRGRI 435
Qy 334 ----- 333
Db 436 LHNAYSLTLOGLISWLGSLRLSRLGSLALIHNNTHLCFVHTVPWDLFRNPQALLH 495
Qy 334 ----- 333
Db 496 TANRPEDECYVGLACHQLCARGHCWGPGPTQCVCNCSQFLRGQCEBECRVLOGLPREYV 555
Qy 334 ----- 333
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
Qy 334 ----- 333
Db 616 FPDEGACQPCPNCTHSCVDLDDKGPABORASPLTSIVSADVGLLVVVLGVVFGILI 675
Qy 334 ----- 333
Db 676 KRRQKIRKYTMRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVILGSAFCTVY 735
Qy 334 ----- 333
Db 736 KGIWPDGENVKIPVAIKVLRNTSPKANKEILDEAYVMAGVGSYVSRLLIGICTSTVQ 795
Qy 334 ----- 333
Db 796 LVTQMPYGCLLDHRNENRGLSGDILLNMCQIAKMSYLEVDVLRVHRDLAARNVLKS 855
Qy 334 ----- 333
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIMWMALESILRRRFTHQSDVMSYGVTVWEL 915
Qy 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVWKMWIDSECRPRELVS 975
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 76.7646 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-4

Perfect score: 3783

Sequence: 1 MRAAPLLARASLSGLF.....CWKPVQKAPPPPAHHHHH 697

Scoring table: BLOSUM62 , Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003s.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3783	100.0	697	4	AAE13111 Human HER
2	3473.5	91.8	690	4	AAE13109 Human HER
3	3014	79.7	564	4	AAE13110 Human HER
4	2957.5	78.2	555	4	AAE13108 Human HER
5	2585	68.3	919	3	AB21203 Human HER
6	2585	68.3	919	5	AA51148 Human HER
7	2478	65.5	479	4	AAE13112 Human HER
8	2422	64.0	1255	3	AA92620 Human her
9	2422	64.0	1255	4	AAE12130 Human tyr
10	2422	64.0	1255	4	AAE12130 Human tyr
11	2422	64.0	1255	5	AAE26349 Human HER
12	2422	64.0	1255	5	AAE26366 Human HER
13	2422	64.0	1255	5	AAU74545 Human HER
14	2422	64.0	1255	6	ABR47407 Breast ca
15	2422	64.0	1255	6	ABP74708 Human HER
16	2422	64.0	1255	6	AAE38390 Human C-e
17	2422	64.0	1255	6	ADA38143 Human erb
18	2422	64.0	1255	7	ADA37255 Human Erb
19	2422	64.0	1255	7	ADB67621 Human epi
20	2420	64.0	1433	2	AA939568 Sequence
21	2417	63.9	1255	2	AAW01111 HER-2/neu
22	2417	63.9	1255	2	AAW92406 Human HER
23	2417	63.9	1255	3	AA984780 Amino aci
24	2417	63.9	1255	3	AA211198 Human HER
25	2417	63.9	1255	4	AA988267 HER2/neu

ALIGNMENTS

RESULT 1

AAE13111
ID AAE13111 standard; protein; 697 AA.

AC AAE13111;

XX
XX
DT 28-JAN-2002 (first entry)

XX Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.

KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

XX Homo sapiens.

OS Rattus norvegicus.

OS Unidentified.

OS Chimeric.

XX WO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010515.

XX 30-MAR-2000; 2000US-0193504P.

XX (DEND-) DENDREON CORP.

XX Laus R, Vidovic D, Graddis T;

XX WPI; 2001-662965/76.

XX N-PSDB; AAD21567.

XX An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

XX Claim 7; Page 27; 59pp; English.

XX The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to

26	2417	63.9	1255	4	AA885458	Human HER
27	2417	63.9	1255	5	AAE20479	Human Her
28	2417	63.9	1255	5	AAU77114	Human Her
29	2417	63.9	1255	5	AA51143	Human Her
30	2417	63.9	1255	5	AAE24067	Human Her
31	2417	63.9	1255	6	ABR43687	Human C-e
32	2417	63.9	1255	7	ABR82066	Human Her
33	2417	63.9	1255	7	ADC09593	Her2/Neu
34	2417	63.9	1255	7	ADD25484	Binding d
35	2417	63.9	1255	7	ADE63281	Human Pro
36	2417	63.9	1255	7	ADE76190	Human HER
37	2410	63.7	1253	7	ADC35106	Human bre
38	2346	62.0	1223	5	AAU98923	Human bre
39	2142.5	56.6	920	5	AA51152	Mouse Her
40	2142.5	56.6	926	5	AA51153	Mouse Her
41	2104	55.6	1200	3	AA21208	Human HER
42	2001.5	52.9	782	2	AAW19764	Her2-GM-C
43	1974.5	52.2	1256	3	AA21206	Mouse Her
44	1974.5	52.2	1256	4	AA62860	Amino aci
45	1974.5	52.2	1256	5	AA51151	Mouse Her

CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFF or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular domain, an Ala
CC linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2
CC membrane distal intracellular domain, an Ala Ala linker, a mature rat
CC granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a
CC C-terminal tag
XX
SQ Sequence 697 AA;

Query Match 100.0%; Score 3783; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.2e-243;
Matches 697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAAPLLIARAASISLGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60
Db 1 MRAAPLLIARAASISLGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQGVVQGNLELTPTNASLSFLQDIEQVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLQYQGVVQGNLELTPTNASLSFLQDIEQVQGVLIHNVQVPLQRLRI 120

Qy 121 VRGTQLPEDNYALAVLQNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQL 180
Db 121 VRGTQLPEDNYALAVLQNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLETDCHECAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTPES 300
Db 241 GGCARCKGPLETDCHECAAGCTGPKHSDCLACLHFNHSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGRYTFGASCVTACPNYVSTVDSASIIINPEKLGAGGMVHRRSSSTRSGGDL 360
Db 301 MPNPEGRYTFGASCVTACPNYVSTVDSASIIINPEKLGAGGMVHRRSSSTRSGGDL 360

Qy 361 TLGLEPSEEAAPRSLAPSEAGSDVFDGLGMGAAGLQSLPHTDPSFLQRYSEDPTVP 420
Db 361 TLGLEPSEEAAPRSLAPSEAGSDVFDGLGMGAAGLQSLPHTDPSFLQRYSEDPTVP 420

Qy 421 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAATLSPGKNG 480
Db 421 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAATLSPGKNG 480

Qy 481 VVKDVAFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPERGAPPSTFKG 540
Db 481 VVKDVAFAGGAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQDPPERGAPPSTFKG 540

Qy 541 TPTAENPYGLDVPAAAPTPSPNPTVTRPKHVDIAKEALSLNDMRALENEKNEEDVDII 600
Db 541 TPTAENPYGLDVPAAAPTPSPNPTVTRPKHVDIAKEALSLNDMRALENEKNEEDVDII 600

Qy 601 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTDCEIEVTF 660
Db 601 SNEFSIORPTCVOTRLKLYKQGLRGNLTKLNGALTMIAHYQTNCPPTDCEIEVTF 660

Qy 661 EDFIKNLKGLFDIPDFDCKVPQKGAAPPPPAHHHHH 697
Db 661 EDFIKNLKGLFDIPDFDCKVPQKGAAPPPPAHHHHH 697

RESULT 2
AAE13109
ID AAE13109 standard; protein; 690 AA.
XX
AC AAE13109;
XX

DT 28-JAN-2002 (first entry)
XX Human HER500-hGM-CSF fusion protein construct.
DE Immunostimulatory fusion protein; IFF; antigen component; therapy;
XX immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hGM-CSF fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX WO200174855-A2.
FN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US010515.
XX 30-MAR-2000; 2000US-0193504P.
PR (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
XX WPT; 2001-662965/76.
DR N-PSDB; AAD21565.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
PT of HER-2 and an antigen elicits an immune response to the antigen and is
PT useful for the treatment of associated cancer associated.
XX
XX Claim 7; Page 26; 59pp; English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
CC nucleic acid molecules encoding such proteins. The IFPs comprise a
CC polypeptide antigen component and an immunostimulatory component derived
CC from the intracellular domain of HER-2 protein which is effective to
CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC immune response to the antigen. IFF or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC hGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular and intracellular
CC domains, an Ala Ala linker, a mature human granulocyte- macrophage colony
CC stimulating factor (GM-CSF) sequence and a C-terminal tag
XX
SQ Sequence 690 AA;

Query Match 91.8%; Score 3473.5; DB 4; Length 690;
Best Local Similarity 91.7%; Pred. No. 2.3e-222;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

Qy 1 MRAAPLLIARAASISLGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60
Db 1 MRAAPLLIARAASISLGLFLFLFWLDRSVLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQGVVQGNLELTPTNASLSFLQDIEQVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLQYQGVVQGNLELTPTNASLSFLQDIEQVQGVLIHNVQVPLQRLRI 120

Qy 121 VRGTQLPEDNYALAVLQNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQL 180
Db 121 VRGTQLPEDNYALAVLQNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGVLIQRNPQL 180

Qy 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDNRSRACHPCSPMCKGRCWGESSEDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 3803.06 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARASLSGLFLP.....ENPEYLGLDVPAAHHHHH 564

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+p2n.model
-Q=/cgn2_1/USPTC_spool/US9821883/runat_09092004_105127_6643/app_query.fasta_1.3100
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -IOCPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.coi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US9821883 @CGN 1 1.1571 @runat_09092004_105127_6643 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gsa_hum.*
18: em_gsa_inv.*
19: em_gsa_pln.*
20: em_gsa_vrt.*
21: em_gsa_fun.*
22: em_gsa_nam.*
23: em_gsa_mus.*
24: em_gsa_pro.*
25: em_gsa_rod.*
26: em_gsa_phg.*
27: em_gsa_vrl.*
28: gb_gsa1.*

29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1977.5	64.6	4323	11 AK031099	AK031099 Mus muscu
2	1367	64.2	4463	11 AK083669	AK083669 Mus muscu
3	1291.5	42.2	1016	12 BM562913	BM562913 AGENCOURT
4	1169	38.2	1129	12 BM802792	BM802792 AGENCOURT
5	1157	37.8	871	14 CA488274	CA488274 AGENCOURT
6	1157	37.8	885	14 CA455074	CA455074 AGENCOURT
7	1150	37.6	4715	11 AF318349	AF318349 Homo sapi
8	1131.5	37.0	893	14 CA455141	CA455141 AGENCOURT
9	1084	35.4	1004	14 CD515356	CD515356 AGENCOURT
10	1081	35.3	894	14 CA454131	CA454131 AGENCOURT
11	1080	35.3	653	12 BM721340	BM721340 UI-E-E01-
12	1073	35.0	588	13 BX478931	BX478931 DKFZp686A
13	1072	35.0	902	14 CA488868	CA488868 AGENCOURT
14	1046	34.2	569	9 AL701765	AL701765 DKFZp686P
15	1033	33.7	583	13 BX479114	BX479114 DKFZp686P
16	1022.5	33.4	795	14 CA328613	CA328613 UI-M-FY0-
17	1021	33.3	3110	11 AK031542	AK031542 Mus muscu
18	1010	33.0	808	14 CA489534	CA489534 AGENCOURT
19	991	32.4	616	14 CD721801	CD721801 o101f07.y
20	973	31.8	1988	11 BC023725	BC023725 Mus muscu
21	972	31.7	943	13 BC046553	BC046553 Mus muscu
22	972	31.7	813	13 BQ958632	BQ958632 AGENCOURT
23	931	30.4	813	12 BI557797	BI557797 603236977
24	926	30.2	855	13 BU594980	BU594980 AGENCOURT
25	921	30.1	614	10 AK370693	AK370693 QV1-BT026
26	910	29.7	717	14 CB598701	CB598701 AGENCOURT
27	909.5	29.7	852	14 CD516283	CD516283 AGENCOURT
28	908	29.7	791	12 BI154872	BI154872 602902857
29	901	29.4	491	12 BM790293	BM790293 K-EST0070
30	886	28.9	906	14 CA454570	CA454570 AGENCOURT
31	883	28.8	998	12 BI649877	BI649877 603296516
32	875	28.6	685	9 AU123871	AU123871 AU123871
33	873.5	28.5	932	14 CA487981	CA487981 AGENCOURT
34	868	28.3	609	10 AW701942	AW701942 uq93b02.y
35	862	28.2	621	12 BG283493	BG283493 602407782
36	855	27.9	786	12 BI155788	BI155788 602904360
37	846	27.6	887	13 BQ717097	BQ717097 AGENCOURT
38	835.5	27.3	649	9 AA496412	AA496412 zv37802.r
39	817	26.7	533	14 CB437620	CB437620 685491.MA
40	817	26.7	552	12 BG277542	BG277542 ux45d07.y
41	817	26.7	670	9 AI906012	AI906012 RC-BT105-
42	808	26.4	876	13 BQ769889	BQ769889 UI-M-F10-
43	805	26.3	514	10 BF998814	BF998814 QVO-GN014
44	800	26.1	887	14 CA980253	CA980253 AGENCOURT
45	797.5	26.0	678	12 BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1

AK031099

LOCUS

DEFINITION

AK031099

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AK031099 4323 bp mRNA linear HTC 18-SEP-2003
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:5930404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
AK031099
AK031099.1 GI:26082143
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4

The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 695-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

FEATURES
source

10349636

11042159

11076861

1757-1771 (2000)

695-690 (2001)

563-573 (2002)

bases 1 to 4323

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,

Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,

Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,

Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers

1. .4323

/db_xref="PANTOM_DB:5930404N10"

/db_xref="MGI:2394277"

/db_xref="taxon:10090"

/clone="5930404N10"

/tissue_type="forelimb"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="13 days embryo"

48. .3818

/note="putative

v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,

neuro/glioblastoma derived oncogene homolog (avian)

(MGI:95410, GB|U71126, evidence: BLASTN, 99%,

match=449)"

misc_feature

Alignment Scores:

Pred. No.: 2,43e-114 Length: 4323

Score: 1977.50 Matches: 439

Percent Similarity: 36.72% Conservative: 31

Best Local Similarity: 34.30% Mismatches: 79

Query Match: 64.58% Indels: 731

DB: 11 Gaps: 4

US-09-821-883-3 (1-564) x AK031099 (1-4323)

Qy 3 AlaAlaProLeuLeuAlaAlaAlaSerLeuSerLeuGlyPheLeuLeu 22

Db 33 GCCGAGTAATCATCATGGAGCTGGCGCTGGTGGGGTCTCTCTCGCCCTC 92

Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42

Db 93 -----CTGTCCCGGAGCGCGGTACC 116

Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62

Db 117 CAATGTGTATCCGGTACCGATGAGTGGAGTCCCTGCGAGTCTCGAGACCCACCTG 176

Qy 63 AspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGluLeuThr 82

Db 177 GACATGCTTCCGACCTTACACGGCTGTGAGTGGTGGAGGCAATTGGAGCTTACC 236

Qy 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGlnValGlnGlyTyr 102

Db 237 TACCTGCGCGCAATGCCAGCTCTCATTCCTGCAGGACATCCAGGAAGTCCAGGGATAC 296

Qy 103 ValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgGlyValArg 122

Db 297 ATGCTCATCGCTCACACCGAGTGAACACACCTCCCTGCGAGGTTCCGATCGTGAGA 356

Qy 123 GlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142

Db 357 GGGACTCAGCTCTTTGAGGACAAAGTATGCTGCTGTGTAGACAACCGAGACCTTTG 416

Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161

Db 417 GACAACTGTCACACCGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 476

Qy 162 ArgSerLeuThrGluLeuLeuGlyValLeuIleGlnArgAsnProGlnLeuCys 181

Db 477 CGAAGTCTCAGAGATCTTGAAGGAGGAGTGTGATCCGTGGGAAACCTCAGCTCTGC 536

Qy 182 TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnGlnLeuAlaLeuThr 201

Db 537 TACAGGACATGGTGTGGAGAGATGCTCCGTAAGATATACACAGTGGCTCTCTC 596

Qy 202 LeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221

Db 597 GACATGGACACCAATCGTTCGCGGCTGTCCACCTGTGTGGCCCAACCTGCAAGACA 656

Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrValCysAlaGly 241

Db 657 CACTGTTGGGGTGAGAGTCTCTGAAGAGATGTCTGACTTGGACCATCTGTACTAGT 716

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 681.931 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlh

-MODEL=frame+p2n model
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app.query.fasta_1.3100
-DB=Published Applications NA -Qfmt=fastap -SUFFIX=nrnp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USERS=US09821883 @CNC 1 1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

RESULT 1
US-09-821-883-8
; Sequence 8, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct

ALIGNMENTS

1	3062	100.0	1692	9	US-09-821-883-8	Sequence 8, Appl
2	3014	98.4	2091	9	US-09-821-883-9	Sequence 9, Appl
3	3005.5	96.2	1665	9	US-09-821-883-6	Sequence 6, Appl
4	2957.5	96.6	2070	9	US-09-821-883-7	Sequence 7, Appl
5	2430.5	79.4	4543	9	US-09-769-508-1	Sequence 1, Appl
6	2427.5	79.3	4530	9	US-09-877-177-11	Sequence 11, Appl
7	2427.5	79.3	4530	13	US-10-392-113-45	Sequence 45, Appl
8	2427.5	79.3	4530	13	US-10-426-836-11	Sequence 11, Appl
9	2427.5	79.3	4530	15	US-10-177-293-125	Sequence 125, App
10	2427.5	79.3	4530	15	US-10-007-926A-119	Sequence 119, App
11	2427.5	79.3	4530	15	US-10-338-730-1	Sequence 1, Appl
12	2427.5	79.3	4530	15	US-10-101-510-124	Sequence 124, App
13	2427.5	79.3	4530	16	US-10-116-275-131	Sequence 131, App
14	2427.5	79.3	4530	16	US-10-272-437A-27	Sequence 27, Appl
15	2427.5	79.3	4530	16	US-10-117-937-595	Sequence 595, App
16	2427.5	79.3	4530	16	US-10-159-563-208	Sequence 208, App
17	2427.5	79.3	4530	16	US-10-435-696-10	Sequence 10, Appl
18	2427.5	79.3	4530	17	US-10-734-564-59	Sequence 59, Appl
19	2427.5	79.3	4642	15	US-10-198-846-10896	Sequence 10896, A
20	2425	79.2	9274	9	US-09-811-123-7	Sequence 7, Appl
21	2425	79.2	9274	9	US-09-811-115-1	Sequence 1, Appl
22	2422.5	79.1	4473	10	US-09-441-411-5	Sequence 5, Appl
23	2422.5	79.1	4473	15	US-10-146-473-32	Sequence 32, Appl
24	2422.5	79.1	4473	15	US-10-207-655-44	Sequence 44, Appl
25	2422.5	79.1	4473	15	US-10-101-510-81	Sequence 81, Appl
26	2422	79.1	3765	15	US-10-207-498-5	Sequence 5, Appl
27	2422	79.1	3768	9	US-09-811-123-8	Sequence 8, Appl
28	2422	79.1	3768	9	US-09-811-115-2	Sequence 2, Appl
29	2422	79.1	3768	13	US-09-984-092-3	Sequence 3, Appl
30	2422	79.1	3768	13	US-10-280-576-3	Sequence 3, Appl
31	2422	79.1	3768	17	US-10-441-779C-3	Sequence 3, Appl
32	2417	78.9	3768	9	US-09-854-356-9	Sequence 9, Appl
33	2417	78.9	3768	9	US-09-930-125-1	Sequence 1, Appl
34	2417	78.9	3768	13	US-09-765-373-1	Sequence 1, Appl
35	2417	78.9	3768	15	US-10-313-644-1	Sequence 1, Appl
36	2403.5	78.5	4606	10	US-09-971-392-70	Sequence 70, Appl
37	1974.5	64.5	3771	9	US-09-854-356-11	Sequence 11, Appl
38	1968	64.3	3955	9	US-09-870-759-117	Sequence 117, App
39	1968	64.3	3955	9	US-09-854-356-10	Sequence 10, Appl
40	1968	64.3	3955	10	US-09-751-708A-117	Sequence 117, App
41	1838.5	60.0	1437	9	US-09-821-883-10	Sequence 10, Appl
42	1644	53.7	2132	16	US-10-412-804A-3	Sequence 3, Appl
43	1641.5	53.6	2149	16	US-10-412-804A-9	Sequence 9, Appl
44	1638.5	53.5	2164	16	US-10-412-804A-5	Sequence 5, Appl
45	1587	51.8	867	9	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-8

Alignment Scores:

Pred. No.: 1,22e-282 Length: 1692
Score: 3062.00 Matches: 564
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-821-883-3 (1-564) x US-09-821-883-8 (1-1692)

Qy 1 MetArgAlaAlaProLeuLeuLeuAlaAlaArgAlaAlaSerLeuLeuLeuPhe 20
Db 1 ATGAGAGCTGACCCCTCTCTGGCCAGGAGCAGCAGCCCTTAGCCTTGGCTTCTTGT 60
Qy 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTACTAGCCAAAGAGATTGGCGCGCGGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGGTGCACCGCACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACTGGACATGCTCCGCGCACTCTACCAAGGCTGCCAGGTGGTGCAGGGAACCTGGAA 240
Qy 81 LeuThrTrpLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTCACCTACCTGCCCAACATGCCAGCTGCTCTCTCTGCGAGGATATCCAGGAGGTGCAG 300
Qy 101 GlyTrpValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATGCTGCACCAACAAAGTAGGAGGCTCCCACTGCAGAGGCTCGCGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTrpAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCTGGCGCTGCTAGACAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAAACAATACCAACCTGTACAGGGGCTCCCAAGGAGGCTCGCGGAGGTGCAG 480
Qy 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTGGAAGCCTCACAGAGATCTTGAAGAGAGGGGTCTTGATCCAGCGGAAACCCCAAGCTC 540
Qy 181 CysTrpGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
Db 541 TGCTACCAAGGACACGATTTTGTGGAGAGACATCTTCCACAGNACAACACAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACACGTATAGACACCAACCGCTCTCGGGCTGCCACCCCTGTTCTCCGATGTGTAAGGSC 660
Qy 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTCTCGGGAGAGAGTTCTGAGGATTTGTAGAGCCTGACGCGCACTGTCTGTGCC 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGGCTGCAAGGGGCACTGCCCACTGACTGTGCTGCTGCTGCTGCTGCTGCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTCAGCGGGCCCCAAGCACTCTGACTGCTGGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTrpAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGCACTGCCAGCCCTGGTGCATCTACACACAGACACGTTTGAGTCC 900
Qy 301 MetProAsnProGluGlyArgTrpThrPheGlyAlaSerCysValThrAlaCysProTrp 320

Db 901 ATGCCCAATCCGAGGCGCGGTATACATTCCGCCACAGCTGTGTGACTGCTGTCCCTAC 960
Qy 321 AsnTrpLeuSerThrAspValGlySerAlaSerIleIleAsnPheGluLysLeuGlyAla 340
Db 961 AACTACCTTTTCTACGACGTGGGATCCGCTAGCATCATTAATTTTCGAGAAGTTGGGCGCT 1020
Qy 341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 360
Db 1021 GGGGCGATGTCTCCACACAGCAGCCGAGCTCATCTACAGGAGTGGCGGTGGGACCTG 1080
Qy 361 ThrLeuGlyLeuGluProSerGluGluGluAlaProArgSerProLeuAlaProSerGlu 380
Db 1081 ACACGTAGGCTGGAGCCCTCTGAAGAGAGAGCCGCCAGGTCTCCACTGGCACCCCTCGAA 1140
Qy 381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
Db 1141 GGGGCTGGCTCCGATGTATTGATGGTGACCTGGGAATGGGGGCGCCAGAGGGGCTGCA 1200
Qy 401 SerLeuProThrHisAspProSerProLeuGlnArgTrpSerGluAspProThrValPro 420
Db 1201 AGCTTCCCAACATGACCCAGCCCTCTACAGCGGTACAGTAGGACCCCCACAGTAGTACC 1260
Qy 421 LeuProSerGluThrAspGlyTrpValAlaProLeuThrCysSerProGlnProGluTrp 440
Db 1261 CTGCCCTCTGAGACTGATGGCTAGCTTGGCCCTGACCTGCAGCCCCCAGCCCTGAATAT 1320
Qy 441 ValAsnGlnProAspValArgProGlnProProSerProArgGluGlyProLeuProAla 460
Db 1321 GTGAACCAAGCAGATGTTCGGCCCCCAGCCCCCTTCGCCCGCAGAGGGGCTCTGCCCTGCT 1380
Qy 461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
Db 1381 GCGGACCTGCTGGTGCCACTCTGGAAGGGGCCAAGACTCTCTCCCGAGGGAAGATGGG 1440
Qy 481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTrpLeuThrPro 500
Db 1441 GTCTCAAGACGTTTTTGGCTTTGGGGGTGCCCTGGAGAACCCCGAGTACTTCACACCC 1500
Qy 501 GlnGlyGlyAlaProGlnProHisProProAlaPheSerProAlaPheAspAsn 520
Db 1501 CAGGAGAGAGTGGCTTCCCTCAGCCCCCAGCTCTCTCTGCTTCCAGCCAGCTTCGACAC 1560
Qy 521 LeuTrpTrpAspGlnAspProProGluArgGlyAlaProProSerThrPheLysGly 540
Db 1561 CTCTATTACTGGGACAGGACCCACAGAGCGGGGGCTCCACCCAGCACCTTCAAGGG 1620
Qy 541 ThrProThrAlaGluAsnProGluTrpLeuGlyLeuAspValProAlaAlaAlaHisHis 560
Db 1621 ACACCTACGGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGCGCGCCGACATCAC 1680
Qy 561 HisHisHisHis 564
Db 1681 CATCACCATCAC 1692

RESULT 2

US-09-821-883-9
; Sequence 9, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 97.2259 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLF.....ENPEYGLDVPAHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09821883/runat_09092004.105128.6657/app_query.fasta_1.3100
-DB=Issued_Patents_NA -QMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CIGN 1.1 215 @runat_09092004.105128.6657 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2427.5	79.3	4530	1 US-08-229-515A-9	Sequence 9, Appli
2	2427.5	79.3	4530	1 US-08-645-865-9	Sequence 9, Appli
3	2427.5	79.3	4530	4 US-09-167-322-4	Sequence 4, Appli
4	2427.5	79.3	4530	4 US-09-527-487-1	Sequence 1, Appli
5	2427.5	79.3	4530	4 US-09-877-177A-11	Sequence 11, Appli
6	2425	79.2	9274	4 US-09-811-115-1	Sequence 1, Appli
7	2422.5	79.1	4473	2 US-09-048-804-1	Sequence 1, Appli
8	2422.5	79.1	4473	3 US-09-056-105-26	Sequence 26, Appli
9	2422.5	79.1	4473	4 US-09-663-834A-3	Sequence 3, Appli
10	2422	79.1	3768	4 US-09-811-115-2	Sequence 2, Appli
11	2417	78.9	3768	2 US-08-625-101-1	Sequence 1, Appli
12	2417	78.9	3768	2 US-08-356-786-1	Sequence 1, Appli

13	1968	64.3	3955	1	US-08-229-515A-14	Sequence 14, Appli
14	1968	64.3	3955	1	US-08-645-865-14	Sequence 14, Appli
15	1641.5	53.6	2385	2	US-09-146-283-3	Sequence 3, Appli
16	1641.5	53.6	2385	3	US-08-579-823A-3	Sequence 3, Appli
17	1641.5	53.6	2385	3	US-09-344-195-3	Sequence 3, Appli
18	1626.5	53.1	1872	3	US-08-422-108-2	Sequence 2, Appli
19	1626.5	53.1	1872	4	US-08-422-734-2	Sequence 2, Appli
20	732	23.9	1868	1	US-08-658-883B-1	Sequence 1, Appli
21	732	23.9	1868	4	US-09-676-610B-26	Sequence 26, Appli
22	731.5	23.9	1593	4	US-09-676-610B-25	Sequence 25, Appli
23	731.5	23.9	5532	2	US-08-475-035-3	Sequence 3, Appli
24	731.5	23.9	5532	2	US-09-676-610B-17	Sequence 17, Appli
25	723.5	23.6	5484	3	US-09-632-580A-3	Sequence 3, Appli
26	723.5	23.6	5551	1	US-08-484-438-1	Sequence 1, Appli
27	723.5	23.6	5551	1	US-08-484-438-3	Sequence 3, Appli
28	711	23.2	4905	1	US-07-978-895-3	Sequence 3, Appli
29	711	23.2	4905	1	US-08-473-119-3	Sequence 3, Appli
30	711	23.2	4905	2	US-08-475-352-3	Sequence 3, Appli
31	711	23.2	4905	4	US-09-170-639-3	Sequence 3, Appli
32	711	23.2	4975	3	US-09-630-706-3	Sequence 3, Appli
33	703	23.0	1958	4	US-09-570-454-1	Sequence 1, Appli
34	703	23.0	1958	4	US-09-867-521-1	Sequence 1, Appli
35	624.5	20.4	4545	6	5183884-3	Patent No. 5183884
36	493	16.1	322	1	US-08-421-356-1	Sequence 1, Appli
37	493	16.1	322	4	US-09-046-783-1	Sequence 1, Appli
38	259	8.5	4149	2	US-08-737-715-1	Sequence 1, Appli
39	259	8.5	4723	4	US-09-023-655-1137	Sequence 1137, Ap
40	255.5	8.3	2533	3	US-09-149-922-5	Sequence 5, Appli
41	242	7.9	4975	2	US-08-249-687C-1	Sequence 1, Appli
42	242	7.9	4989	2	US-08-666-392A-3	Sequence 3, Appli
43	242	7.9	4989	2	US-08-625-819-1	Sequence 1, Appli
44	242	7.9	4989	3	US-08-755-558-4	Sequence 4, Appli
45	242	7.9	4989	3	US-08-746-559A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518985
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-9

```

Alignment Scores:

Pred. No.:	1,869-189	Length:	4530
Score:	2427.50	Matches:	522
Percent Similarity:	41.61%	Conservative:	9
Best Local Similarity:	40.91%	Mismatches:	20
Query Match:	79.28%	Indels:	725
DB:	1	Gaps:	4

US-09-821-883-3 (1-564) x US-08-229-515A-9 (1-4530)

Qy	2	ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly	17
Db	94	CGCGCCCTCCACGCGGGTCCCAAGCCGGAGCCATGGGCGCGAGCCGCGATGAGCACCCATG	153
Qy	18	PheLeuPheLeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArg	37
Db	154	GAGCTGGCGGGCTTGTCGGCTGG-----GGGTCCTCTCTGCGCCCTCTTGCCCCCC	204
Qy	38	GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer	57
Db	205	GGAGCCGCGAGCACCAAGTGTGCACCGCACAGACATGAAGCTGCGGCTCCTGCGCAGT	264
Qy	58	ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly	77
Db	265	CCCAGAGACCCACCTGGACATGCTCGCGCACCTCTACCAAGGCTGCGAGGTGTCAGGGA	324
Qy	78	AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln	97
Db	325	AACCTGGAACTCACTACTGCCCAACCAATGCCAGCTGTCTTCTCGCAGGATATCCAG	384
Qy	98	GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg	117
Db	385	GAGGTGCGAGGGCTACGTGCTCATCGCTCACAAACAGTGAGGCGAGTCCCATCTGCAGAGG	444
Qy	118	LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp	137
Db	445	CTGCGGATTTGTCGAGGACACCCAGCTCTTTTAGGACAACTATGSCCTGCGCGTGTAGAC	504
Qy	138	AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg	157
Db	505	AATGGAGACCCGCTGAACAAATACCAACCCCTGTCAAGGGGCCCTCCAGAGAGGCTCGGG	564
Qy	158	GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuIleGlnArgAsn	177
Db	565	GAGCTGCGAGCTTCGAAGCCTCACAGAGATCTTTGAAGGAGGGGTCTTGATCCAGCGAAC	624
Qy	178	ProGlnLeuCysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGln	197
Db	625	CCCCAGCTCTGCTACAGGACACGATTTTGTGGAAGGACATCTCCCAAGAACCAACCCAG	684
Qy	198	LeuAlaLeuThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMet	217
Db	685	CTGGCTTCACACTGATAGACACCAACCGCTCTCGGGCCTGCACCCCTGTCTTCGATG	744
Qy	218	CysLysGlySerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThr	237
Db	745	TGTAAGGGCTCCCGCTGCTGGGGAGAGATTCTGAGGATTGTTCAGAGCCTGACGGCAGCT	804
Qy	238	ValCysAlaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu	257
Db	805	GTCTGTGCGGTGGCTGTGTCGCCGTGCAAGGGGGCCACTGCCCACTGCTGCTGCATGAG	864
Qy	258	GlnCysAlaAlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe	277
Db	865	CAGTGTGCTGCGGCTGCACGGGCCCCAAGCACTCTGACTGCTGGCTGGCTGCCCTCCACTTC	924
Qy	278	AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr	297

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 555.448 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARASLSGLFLF.....ENPEYLGLDVPAHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/sgn2_1/USPTO spo01/US09821883/runat_09092004_105126_6623/app_query.fasta_1.3100

-DB=N Geneseq_29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPECL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09821883 @CGN 1.1.1646 @runat_09092004_105126_6623 -NCPU=6 -ICPU=3

-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT_DSPBLOCK=100 -LONGLOG

-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002a:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	3062	100.0	1692	4	AAD21566	Aad21566 Human HER
2	3014	98.4	2091	4	AAD21567	Aad21567 Human HER
3	3005.5	98.2	1665	4	AAD21564	Aad21564 Human HER
4	2957.5	96.6	2070	4	AAD21565	Aad21565 Human HER
5	2427.5	79.3	4530	2	AAT01585	Aat01585 Her-2/neu
6	2427.5	79.3	4530	2	AAT71253	Aat71253 Human HER
7	2427.5	79.3	4530	3	Aaz60815	Aaz60815 Nucleotid
8	2427.5	79.3	4530	4	AAD19731	Aad19731 Human tyr

9	2427.5	79.3	4530	6	ABN85585	Abn85585 Human HER
10	2427.5	79.3	4530	6	ABZ35012	Abz35012 Human gen
11	2427.5	79.3	4530	6	ABV94128	Abv94128 Breast ca
12	2427.5	79.3	4530	6	ABK83918	Abk83918 Human cdv
13	2427.5	79.3	4530	7	ACC50139	Acc50139 Breast ca
14	2427.5	79.3	4530	7	ABQ83856	Abq83856 Human Her
15	2427.5	79.3	4530	8	AAD58073	Aad58073 Human C-e
16	2427.5	79.3	4530	9	ADC09594	Adc09594 Her2/Neu
17	2425	79.2	9274	6	AAD43934	Aad43934 HER-2 tra
18	2425	79.2	9274	6	ABK14057	Abk14057 Human HER
19	2422.5	79.1	4472	3	AA14812	Aa14812 CDNA enco
20	2422.5	79.1	4473	2	ABQ76220	Abq76220 Human tum
21	2422.5	79.1	4473	2	Aaz31071	Aaz31071 HER-2 nuc
22	2422.5	79.1	4473	6	ABZ34969	Abz34969 Human gen
23	2422.5	79.1	4473	6	AAD38904	Aad38904 Human Her
24	2422.5	79.1	4473	7	ACC69999	Acc69999 Human C-e
25	2422.5	79.1	4473	9	ADC35148	Adc35148 Human bre
26	2422.5	79.1	4473	9	ADD25483	Add25483 Binding d
27	2422	79.1	3765	9	ADB67620	Adb67620 Human epl
28	2422	79.1	3768	3	AAA09455	Aaa09455 Human her
29	2422	79.1	3768	6	ABV78168	Abv78168 Human ERB
30	2422	79.1	3768	6	ABZ35744	Abz35744 Human ERB
31	2422	79.1	3768	6	AAD43935	Aad43935 Human HER
32	2422	79.1	3768	6	ABX09987	Abx09987 Human ERB
33	2422	79.1	3768	6	AAD43986	Aad43986 Human HER
34	2422	79.1	3768	6	ABK14058	Abk14058 Human HER
35	2422	79.1	3768	6	ABL91709	Ab191709 Human pol
36	2422	79.1	3768	7	ACC57649	Acc57649 Human pro
37	2418	79.0	9274	4	AAF24297	Aaf24297 HER2 tran
38	2417	78.9	3768	2	AAT40739	Aat40739 HER-2/neu
39	2417	78.9	3768	2	AAx01912	Aax01912 Human HER
40	2417	78.9	3768	5	AAH23392	Aah23392 Human HER
41	2417	78.9	3768	6	AAD32743	Aad32743 Human Her
42	2417	78.9	3768	6	ABK10730	Abk10730 Human Her
43	2417	78.9	3768	6	ABA92250	Ab92250 Human Her
44	2402.5	78.5	4606	9	ABD47370	Abd47370 Human cdn
45	2401	78.4	4299	2	AAQ46083	Aaq46083 Sequence

ALIGNMENTS

RESULT 1

AAD21566

ID AAD21566 standard; DNA; 1692 BP.

XX AC AAD21566;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500 fusion DNA construct comprising OVA-derived octapeptide.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;

KW immunostimulatory component; T-cell mediated immune response; DC;

KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;

KW PAP protein; Ala Arg linker; membrane distal extracellular domain;

KW membrane distal intracellular domain; C-terminal tag; human; OVA;

KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion DNA; ds.

XX OS Homo sapiens.

OS Unidentified.

OS Chimeric.

XX FH Key

XX CDS

XX Location/Qualifiers

FT 1..1692

FT /*tag= a

FT /product= "Human HER500 fusion protein construct

FT comprising human PAP signal sequence, mature PAP protein,

FT an Ala Arg linker, human HER-2 signal sequence, mature

FT HER-2 membrane distal extracellular domain, an Ala

FT linker, an ovalbumin (OVA)-derived immunodominant

FT octapeptide, HER-2 membrane distal intracellular domain

FT and a C-terminal tag"

FT /note= "CDS does not include stop codon"

```
FT /partial
XX WO200174855-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010515.
XX
XX 30-MAR-2000; 2000US-0193504P.
XX
XX (DND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
XX
XX WPI; 2001-662965/76.
XX
XX P-PSDB; AAE13110.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
XX of HER-2 and an antigen elicits an immune response to the antigen and is
XX useful for the treatment of associated cancer associated.
XX
XX Example 3; Page 28; 59pp; English.
XX
XX The invention relates to immunostimulatory fusion proteins (IFP) and
XX nucleic acid molecules encoding such proteins. The IFPs comprise a
XX polypeptide antigen component and an immunostimulatory component derived
XX from the intracellular domain of HER-2 protein which is effective to
XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX immune response to the antigen. IFP or superactivated dendritic cells are
XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX associated with a particularly antigen. The present sequence is HER500
XX fusion DNA construct which comprises DNA molecules encoding human PAP
XX signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
XX signal sequence, mature HER-2 membrane distal extracellular domain, an
XX Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2
XX membrane distal intracellular domain and a C-terminal tag
XX
XX Sequence 1692 BP; 338 A; 573 C; 470 G; 311 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3.3e-146 Length: 1692
XX Score: 3062.00 Matches: 564
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 4 Gaps: 0
XX
XX US-09-821-883-3 (1-564) x AAD21566 (1-1692)
XX
XX 1 MetArgAlaAlaProLeuLeuLeuAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
XX
XX 1 ATGAGAGTGACCCCTCTCTGCGCAGGGCAGCAGCCCTAGCCTTGGCTTCTTGT 60
XX
XX 21 LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAla 40
XX
XX 61 CTGCTTTTTTCTGCTAGACCGAAGTGTTACTAGCCAAAGAGTGTGGCGCGGGCGCG 120
XX
XX 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
XX
XX 121 TCGACCCCAAGTGTGCACCGGCACACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACC 180
XX
XX 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
XX
XX 181 CACCTGGACATGCTCCGCCACTCTTACAGGGCTGCCAGGTGGTGACAGGAACTGGAA 240
XX
XX 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
XX
XX 241 CTCACCTACCTGCCCAACCAATGCCAGCTGTCTTCTCTGCAGATATCCAGGAGGTGCAG 300
XX
XX 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
XX
XX 301 GGCTACGTGCTCATCGCTCACAAACCAAGTAGGGCAGGTCCCACTGCAGAGGTGCGGATT 360
XX
XX
XX 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
XX
XX 361 GTGCGAGGCACCCAGCTCTTTTGGAGACAATATATGCTGGCCGTGTAGACAATGGAGAC 420
XX
XX 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
XX
XX 421 CCGCTGAACAATAACACCCCTGTACAGGGGCCCTCCCAAGAGAGGCTGCGGAGCTGCAG 480
XX
XX 161 LeuArgSerLeuThrGluIleLeuLysGlyValLeuLeuGlnArgAsnProGlnLeu 180
XX
XX 481 CTTCGAAGCCTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGGAACCCCGAGCTC 540
XX
XX 181 CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
XX
XX 541 TGCTACCAGGACACGATTTTGTGAAGACATCTTCCACAAGAAACAACACCTGGCTCTC 600
XX
XX 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
XX
XX 601 ACCTGATAGACACCAACCGCTCTCGGGCCCTGCGACCCCTGTTTCCGATGTGTAAGGCG 660
XX
XX 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
XX
XX 661 TCCCGCTGCTGGGAGAGAGTTCTGAGATTGTTCAGAGCCTGACGCGCACTGTCTGTGCC 720
XX
XX 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlnGlnCysAla 260
XX
XX 721 GGTGGCTGTGCCCTGCAAGGGGCCACTGCCCCACTGACTGCTGCATGACGAGCTGTCT 780
XX
XX 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
XX
XX 781 GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCCCTCCACTTCAACCAAGT 840
XX
XX 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
XX
XX 841 GGCACTGTGAGCTGCATGCCAGCCCTGCTACCTACCAACACAGACACAGCTTTGAGTCC 900
XX
XX 301 MetProLeuProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
XX
XX 901 ATGCCCAATCCGAGGGCCGGTATACATTCGCGCCAGCTGTGTGACTGCTGCTCCCTTAC 960
XX
XX 321 AsnTyrLeuSerThrAspValGlySerAlaSerIleLeuAsnPheGluLysLeuGlyAla 340
XX
XX 961 AACTACCTCTTCTACGGAGCTGGGATCCCTAGCATCATTAATTTTCGAGAAGTTGGCGCT 1020
XX
XX 341 GlyGlyMetValHisHisArgHisArgSerSerThrArgSerGlyGlyGlyAspLeu 360
XX
XX 1021 GGGGCGATGTTCCACCACAGCCAGCCAGCTCATCTACAGAGTGGCGGTGGGACCTG 1080
XX
XX 361 ThrLeuGlyLeuGluProSerGluGluAlaProArgSerProLeuAlaProSerGlu 380
XX
XX 1081 ACACCTAGGGCTGGAGCCCTCTGAAGAGAGAGGCCCCCAAGTCTCCACTGGCAACCTCCGAA 1140
XX
XX 381 GlyAlaGlySerAspValPheAspGlyAspLeuGlyMetGlyAlaAlaLysGlyLeuGln 400
XX
XX 1141 GGGGCTGGCTCCGATGATTTGATGGTACCTGGATGGGGGACAGCCAGGGGCTGCA 1200
XX
XX 401 SerLeuProThrHisAspProSerProLeuGlnArgTyrSerGluAspProThrValPro 420
XX
XX 1201 AGCTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTAGAGACCCACACAGTACC 1260
XX
XX 421 LeuProSerGluThrAspGlyTyrValAlaProLeuThrCysSerProGlnProGluTyr 440
XX
XX 1261 CTGCGCTCTGAGACTGATGGCTAGCTTGGCCCTCTGACTGCGAGCCCGCAGCTGAATAT 1320
XX
XX 441 ValAsnGlnProAspValArgProGlnProProArgGluGlyProLeuProAla 460
XX
XX 1321 GTGAACACAGCAGATGTTTGGGCCCCAGCCCTCTGCCCGCAGAGGGGCCCTCTGCTGCT 1380
XX
XX 461 AlaArgProAlaGlyAlaThrLeuGluArgAlaLysThrLeuSerProGlyLysAsnGly 480
XX
XX 1381 GCCCGACCTGCTGGTGCCACTCTGGAAGGGCCCAAGACTCTCTCCCGAGGGAAGAATGGG 1440
XX
XX 481 ValValLysAspValPheAlaPheGlyGlyAlaValGluAsnProGluTyrLeuThrPro 500
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 6042.41 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLIARAASLSGLFLF.....ENPEYGLDVPAAAHHHHHH 564

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09821883/runat_09092004.105127.6631/app.query.fasta_1.3100
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MAIRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1.1 16795 @runat_09092004.105127.6631 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg.hum:*
31: em.htg.inv:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.pln:*
35: em.htg.rod:*
36: em.htg.mam:*
37: em.htg.vrt:*
38: em.sy:*
39: em.htgo.hum:*
40: em.htgo.mus:*
41: em.htgo.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3062	100.0	1692	6	AX268287 Sequence
2	3014	98.4	2091	6	AX268288 Sequence
3	3005.5	98.2	1665	6	AX268285 Sequence
4	2957.5	96.6	2070	6	AX268286 Sequence
5	2427.5	79.3	4530	6	I21124 Sequence 9
6	2427.5	79.3	4530	6	I59745 Sequence 9
7	2427.5	79.3	4530	6	AR202597 Sequence
8	2427.5	79.3	4530	6	AR283481 Sequence
9	2427.5	79.3	4530	6	AR344811 Sequence
10	2427.5	79.3	4530	6	AX282577 Sequence
11	2427.5	79.3	4530	6	AX587649 Sequence
12	2427.5	79.3	4530	6	AX644071 Sequence
13	2427.5	79.3	4530	6	AX771418 Sequence
14	2427.5	79.3	4530	6	BD005474 Cellular
15	2427.5	79.3	4530	9	HUMHER2A
16	2425	79.2	9274	6	AR409602 Sequence
17	2425	79.2	9274	6	AX060703 Sequence
18	2422.5	79.1	4473	6	AR080259 Sequence
19	2422.5	79.1	4473	6	AR167390 Sequence
20	2422.5	79.1	4473	6	AR392088 Sequence
21	2422.5	79.1	4473	9	HSERB2R
22	2422	79.1	3768	6	AR409603 Sequence
23	2422	79.1	3768	6	AX060704 Sequence
24	2422	79.1	3768	6	AX467229 Sequence
25	2422	79.1	3768	6	AX481438 Sequence
26	2422	79.1	3768	6	BD224136 Novel met
27	2417	78.9	3768	6	AR034479 Sequence
28	2417	78.9	3768	6	BD267514 HER-2/neu
29	2417	78.9	3768	6	AX201817 Sequence
30	2417	78.9	3768	6	AX380923 Sequence
31	2417	78.9	3768	6	AX384604 Sequence
32	2417	78.9	3768	6	AX465456 Sequence
33	2346	76.6	3678	6	AX505114 Sequence
34	2142.5	70.0	2763	6	AX380942 Sequence
35	2142.5	70.0	2781	6	AX380944 Sequence
36	2136	69.8	3780	4	AB008451 Canis fam
37	1990.5	65.0	4506	10	AK129487 Mus muscu
38	1977	64.6	4694	10	BC046811 Mus muscu
39	1977	64.6	4695	10	BC053078 Mus muscu
40	1974.5	64.5	3771	6	BD267516 HER-2/neu
41	1974.5	64.5	3771	6	AX189662 Sequence
42	1974.5	64.5	3771	6	AX380925 Sequence
43	1969.5	64.3	4062	10	HAMNEU
44	1968.5	64.3	4727	10	AY116182 Rattus no
45	1968	64.3	3955	6	BD267515 HER-2/neu

ALIGNMENTS

RESULT 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 56.0399 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLP.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_25:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2136	69.8	1259	6	018735
2	1862.5	64.1	1259	11	08K3F9
3	1612.5	52.7	419	4	Q9UK79
4	1380	45.1	711	11	Q80Y89
5	1014	33.1	881	11	Q8C0E7
6	966	31.5	367	11	Q8R2X1
7	906	29.6	165	4	Q14256
8	894.5	29.2	412	4	Q8WTV0
9	859	28.1	431	13	Q7SY19
10	751	24.5	527	13	Q90836
11	723	23.6	149	6	Q9BG66
12	717.5	23.4	643	11	Q9ERV6
13	717.5	23.4	655	11	Q9WVF5
14	717.5	23.4	1210	11	Q9EP98
15	713.5	23.3	1209	11	Q9GX70
16	708.5	23.1	1209	6	Q8MIL8

17	703	23.0	1191	13	Q7SZF7	Q7szf7 brachydanio
18	701.5	22.9	478	11	Q9ESE0	Q9ese0 rattus norv
19	697	22.8	331	4	Q9BUD7	Q9bud7 homo sapien
20	660	21.6	1165	13	Q9YH40	Q9yh40 xiphophorus
21	658	21.5	1305	13	Q8AW81	Q8aw81 brachydanio
22	635.5	20.8	599	13	Q9PSH2	Q9psh2 gallus gall
23	610	19.9	1328	13	P79754	P79754 fugu rubrip
24	559.5	18.3	1377	5	Q8MLW0	Q8mlw0 drosophila
25	543	17.7	1433	5	Q9BTH9	Q9bth9 anopheles g
26	540.5	17.7	1322	5	Q86NZ2	Q86nz2 drosophila
27	409.5	13.4	1137	13	Q9W6F6	Q9w6f6 gallus gall
28	400.5	13.1	150	6	Q9BG64	Q9bg64 oryctolagus
29	395	12.9	151	6	Q9BG65	Q9bg65 oryctolagus
30	377	12.3	1368	5	Q23821	Q23821 caenorhabdi
31	341	11.1	366	5	Q26569	Q26569 schistosoma
32	341	11.1	1717	5	Q26566	Q26566 schistosoma
33	331	10.8	334	5	Q26567	Q26567 schistosoma
34	331	10.8	342	5	Q26568	Q26568 schistosoma
35	321.5	10.5	1193	5	Q9YLX8	Q9ylx8 ephydatia f
36	305	10.0	1564	5	Q86MD7	Q86md7 echinococcu
37	290.5	9.5	1472	5	Q9USA8	Q9usa8 bombyx mori
38	289	9.4	1671	5	Q9NJV5	Q9njv5 biophalari
39	276.5	9.0	1749	5	Q8TOW6	Q8tow6 echinococcu
40	273.5	8.9	89	11	Q88459	Q88459 mus muscucu
41	273.5	8.9	1418	13	Q8UW83	Q8uw83 paralichthy
42	272.5	8.9	1418	13	Q93457	Q93457 scophthalmu
43	271	8.9	2144	5	Q9VD94	Q9vd94 drosophila
44	263	8.6	1358	13	Q73798	Q73798 xenopus lae
45	262.5	8.6	1369	13	Q8UW86	Q8uw86 paralichthy

ALIGNMENTS

RESULT 1

018735 PRELIMINARY; PRT; 1259 AA.

AC 018735;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Erbb-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "cDNA cloning of erbb-2 from canine mammary gland.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1; -
DR HSSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005506; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:000468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . . ; IEA.
DR InterPro; IPR002048; EGFR_L domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR000494; EGFR_L domain.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 10.3528 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARASLSGLFL.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	78.9	1255	1 ERB2 HUMAN	P04626 homo sapien
2	1959	64.0	1257	1 ERB2 RAT	P06494 rattus norv
3	1946.5	63.6	1254	1 ERB2 MESAU	P06553 mesocricetu
4	752	24.6	703	1 EGFR CHICK	P13387 gallus gall
5	723.5	23.6	1308	1 ERB4 HUMAN	Q15303 homo sapien
6	719.5	23.5	1210	1 EGFR HUMAN	P00533 homo sapien
7	719.5	23.5	1308	1 ERB4 RAT	Q62956 rattus norv
8	717.5	23.4	1210	1 EGFR MOUSE	Q01279 mus musculu
9	697	22.8	1342	1 ERB3 HUMAN	P21860 homo sapien
10	668	21.8	1339	1 ERB3 RAT	Q62799 rattus norv
11	651	21.3	1167	1 XMRK XIPMA	P13388 xiphophorus
12	539.5	17.6	1426	1 EGFR DROME	P04412 drosophila
13	404.5	13.2	1367	1 LT23 CAEL	P24348 caenorhabdi
14	318	10.4	1363	1 ILPR BRALA	Q02466 brachiolesto
15	301.5	9.8	245	1 ERB2 MOUSE	P70424 mus musculu
16	276.5	9.0	1477	1 HTK7 HYDAT	Q25197 hydra atten
17	271	8.9	2146	1 INSR DROME	P09208 drosophila
18	258	8.4	1382	1 INSR HUMAN	P06213 homo sapien
19	252.5	8.2	1372	1 INSR MOUSE	P15208 mus musculu
20	252	8.2	1607	1 MIR LYMT	Q25410 lymnaea sca
21	251.5	8.2	1300	1 IRR MOUSE	Q9wt14 mus musculu
22	249	8.1	1383	1 INSR RAT	P15127 rattus norv
23	248.5	8.1	1297	1 IRR HUMAN	P14616 homo sapien
24	246.5	8.1	1300	1 IRR CAVPO	P14617 cavia porce
25	238.5	7.8	1367	1 IGRF HUMAN	P08069 homo sapien
26	236.5	7.7	1370	1 IGRF RAT	P24062 rattus norv
27	234.5	7.7	1373	1 IGRF MOUSE	Q06751 mus musculu
28	228.5	7.5	581	1 IRR RAT	Q64716 rattus norv
29	207	6.8	1390	1 INSR AEDAE	Q93105 aedes aegypt
30	168	5.5	386	1 PPAP HUMAN	P15309 homo sapien
31	155	5.1	5262	1 MLU2 HUMAN	Q14686 homo sapien
32	153.5	5.0	707	1 SFPO HUMAN	P23246 homo sapien
33	148.5	4.8	1321	1 IRS2 MOUSE	P81122 mus musculu

RESULT 1	34	141.5	4.6	886	1	SMGB_MOUSE	054951	mus musculu
ERB2_HUMAN	35	140	4.6	1696	1	PKCS_BRACL	Q9nj15	branchioosto
ID ERB2_HUMAN	36	139	4.5	634	1	ERBB_ALV	P00534	avian leuko
AC P04626;	37	139	4.5	2442	1	CBP_HUMAN	Q92793	homo sapien
DT 13-AUG-1987 (Rel. 05, Created)	38	138	4.5	830	1	SREB_HUMAN	Q14162	homo sapien
DT 13-AUG-1987 (Rel. 05, Last sequence update)	39	136	4.4	321	1	PRP1_HUMAN	P04380	homo sapien
DT 10-OCT-2003 (Rel. 42, Last annotation update)	40	135.5	4.4	553	1	ODO2_MYCTU	Q10381	mycobacteri
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)	41	135	4.4	620	1	EXTN_TOBAC	P13983	nicotiana t
DE (p18serbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell	42	134.5	4.4	863	1	MLI3_HUMAN	Q8n3f8	homo sapien
DE surface receptor HER2) (MLN 19).	43	132	4.3	833	1	SR2_MOUSE	P59222	mus musculu
GN ERB2 OR HER2 OR NGL OR NEU.	44	130.5	4.3	296	1	CC01_CAEL	P08124	caenorhabdi
OS Homo sapiens (Human).	45	130	4.2	1051	1	ULK1_MOUSE	O70405	mus musculu
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX NCBI_TaxID=9606;								
RN [1]								
RP SEQUENCE FROM N.A.								
RX MEDLINE=86118663; PubMed=3003577;								
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,								
RT "Similarity of protein encoded by the human c-erb-B-2 gene to								
RT epidermal growth factor receptor.";								
RL Nature 319:230-234(1986).								
[2]								
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.								
RX MEDLINE=86070181; PubMed=2999974;								
RA Cousens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,								
RA McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,								
RA Francke U., Levinson A., Ullrich A.;								
RT "Tyrosine kinase receptor with extensive homology to EGF receptor								
RT shares chromosomal location with neu oncogene.";								
RL Science 230:1132-1139(1985).								
[3]								
RN SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.								
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,								
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,								
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;								
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.								
[4]								
RN SEQUENCE OF 737-1031 FROM N.A.								
RX MEDLINE=86016729; PubMed=2995967;								
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;								
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the								
RT c-erbB-1/epidermal growth factor-receptor gene and is amplified in a								
RT human salivary gland adenocarcinoma.";								
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).								
[5]								
RN VARIANTS VAL-654 AND VAL-655								
RP MEDLINE=93194196; PubMed=8095488;								
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;								
RT "Characterization of a new allele of the human ERBB2 gene by allele-								
RT specific competition hybridization.";								
RL Genomics 15:426-429(1993).								
CC -!- FUNCTION: Essential component of a neurogulin-receptor complex,								
CC although neurogulin do not interact with it alone. GP30 is a								
CC potential ligand for this receptor. Not activated by EGF, TGF-								

ALIGNMENTS

CC alpha and amphiregulin.
 CC -|- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -|- SUBUNIT: Heterodimer with each of the other ERBB receptors
 CC (potential). Interacts with PRKCAP (By similarity).
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- PTM: Ligand-binding increases phosphorylation on tyrosine
 CC residues (By similarity).
 CC -|- POLYMORPHISM: There are four alleles due to the variations in
 CC positions 654 and 655. Allele B1 (Ile-654/Val-655) has a frequency
 CC of 0.782; allele B2 (Ile-654/Val-655) has a frequency of 0.206;
 CC allele B3 (Val-654/Val-655) has a frequency of 0.012.
 CC -|- SIMILARITY: Belongs to the EGF receptor family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M11767; AAA35808.1; -.
 CC EMBL; M11761; AAA35808.1; JOINED.
 CC EMBL; M11762; AAA35808.1; JOINED.
 CC EMBL; M11763; AAA35808.1; JOINED.
 CC EMBL; M11764; AAA35808.1; JOINED.
 CC EMBL; M11765; AAA35808.1; JOINED.
 CC EMBL; M11766; AAA35808.1; JOINED.
 CC EMBL; M11730; AAA75493.1; -.
 CC EMBL; M12036; AAA35978.1; -.
 CC EMBL; AX208911; AAQ18082.1; -.
 CC EMBL; X03363; CAA27060.1; -.
 CC FIR; A24571; A24571.
 CC PDB; 1N82; 18-FEB-03.
 CC PDB; 1QR1; 01-JAN-00.
 CC Genew; HGNC:3430; ERBB2.
 CC MIM; 164870; -.
 CC GO; GO:0005012; F.Receptor activity; TAS.
 CC GO; GO:0004716; F.Receptor signaling protein tyrosine kinase . . . ; TAS.
 CC GO; GO:0002833; P.cell proliferation; TAS.
 CC GO; GO:0007167; P.enzyme linked receptor protein signaling pa. . . ; TAS.
 CC GO; GO:0006470; P.protein amino acid dephosphorylation; TAS.
 CC GO; GO:0006468; P.protein amino acid phosphorylation; TAS.
 CC InterPro; IPR000494; EGFR_L domain.
 CC InterPro; IPR006211; Furin-like.
 CC InterPro; IPR006212; Furin repeat.
 CC InterPro; IPR009030; Grow fac recep.
 CC InterPro; IPR000719; Prot kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC InterPro; IPR004019; YLP motif.
 CC Pfam; PF00757; Furin-like; 1.
 CC Pfam; PF00069; pkinase; 1.
 CC Pfam; PF01030; Recep_L domain; 2.
 CC Pfam; PF02757; YLP_2.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot kinase; 1.
 CC SMART; SM00261; FU; 4.
 CC SMART; SM00219; TyrKC; 1.
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane; Glycoprotein; Multigene family; Receptor; Signal;
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW Polymorphism; 3D-structure.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 1255 RECEPTOR PROTEIN-TYROSINE KINASE ERBB-2.
 FT DOMAIN 22 652 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 653 675 POTENTIAL.
 FT DOMAIN 676 1255 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 720 987 PROTEIN KINASE.
 FT NP_BIND 726 734 ATP (BY SIMILARITY).
 FT -----

FT BINDING	753	ATP (BY SIMILARITY).
FT ACT_SITE	845	BY SIMILARITY.
FT DISULFID	195	BY SIMILARITY.
FT DISULFID	199	BY SIMILARITY.
FT DISULFID	220	BY SIMILARITY.
FT DISULFID	224	BY SIMILARITY.
FT DISULFID	236	BY SIMILARITY.
FT DISULFID	240	BY SIMILARITY.
FT DISULFID	255	BY SIMILARITY.
FT DISULFID	288	BY SIMILARITY.
FT DISULFID	299	BY SIMILARITY.
FT DISULFID	315	BY SIMILARITY.
FT DISULFID	334	BY SIMILARITY.
FT DISULFID	511	BY SIMILARITY.
FT DISULFID	515	BY SIMILARITY.
FT DISULFID	531	BY SIMILARITY.
FT DISULFID	544	BY SIMILARITY.
FT DISULFID	563	BY SIMILARITY.
FT DISULFID	567	BY SIMILARITY.
FT DISULFID	587	BY SIMILARITY.
FT DISULFID	600	BY SIMILARITY.
FT DISULFID	626	BY SIMILARITY.
FT DISULFID	630	BY SIMILARITY.
FT MOD_RES	1139	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	1248	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD	68	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	124	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	187	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	259	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	530	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	571	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	629	N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT	452	W -> C.
FT VARIANT	654	/FTID=VAR_016317.
FT VARIANT	655	I -> V (in allele B3; dbSNP:1801201).
FT VARIANT	655	I -> V (in allele B2 and allele B3; dbSNP:1801200).
FT VARIANT	1170	/FTID=VAR_004078.
FT VARIANT	1170	P -> A.
FT SEQUENCE	1255 AA; 137909 MW; 3959DFDA04DCF962 CRC64;	/FTID=VAR_016318.

Query Match 78.9%; Score 2417; DB 1; Length 1255;
 Best Local Similarity 41.2%; Pred. No. 9.8e-138;
 Matches 511; Conservative 4; Mismatches 6; Indels 718; Gaps 2;

QY	35	LARGAASQTQVCTGTDMLRLPASPTHLDMLRHLVGGQVQVQGNLELYLPTNALSFLQ 94
Db	16	LPPGAASQTQVCTGTDMLRLPASPTHLDMLRHLVGGQVQVQGNLELYLPTNALSFLQ 75
QY	95	DIQEVQGYVLIHNNQVQVPLQRLIRVGTQLFEDNYALAVLDNGDPLNTPVTGASG 154
Db	76	DIQEVQGYVLIHNNQVQVPLQRLIRVGTQLFEDNYALAVLDNGDPLNTPVTGASG 135
QY	155	GLRELQLRSLTEILKGGVLIQNPOLCYQDTTLWKDI FHNKNQLALTLDITNRSRACHPC 214
Db	136	GLRELQLRSLTEILKGGVLIQNPOLCYQDTTLWKDI FHNKNQLALTLDITNRSRACHPC 195
QY	215	SPMKGSRGWGESSEDCQSLTRTVCCAGCARCKGLPTDCCHQCAAGCTGPRHSDCLAC 274
Db	196	SPMKGSRGWGESSEDCQSLTRTVCCAGCARCKGLPTDCCHQCAAGCTGPRHSDCLAC 255
QY	275	LHFNHSGICELKCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTVDGNSAII- 333
Db	256	LHFNHSGICELKCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLTVDGNSCTI- 315
QY	334	----- 333
Db	316	PLHNOEVTAEADCTQRCCKSPCARVCYGLGMEHLREVRAVTSANIOBPAGCKKIFGSLA 375
QY	334	----- 333

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 18.0048 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLP.....ENPEYGLDVPAAHHHHH 564
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:
1: piri:
2: pir2:
3: pir3:
4: pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2417	78.9	1255	1 A24571	protein-tyrosine k
2	1962	64.1	1260	1 TVRTNU	protein-tyrosine k
3	1946.5	63.6	1254	2 I48161	p-185 precursor -
4	752	24.6	1223	1 TVCHLV	epidermal growth f
5	751	24.5	527	2 A42032	epidermal growth f
6	723.5	23.6	1308	2 A47253	epidermal growth f
7	719.5	23.5	1210	1 GQHUE	epidermal growth f
8	717.5	23.4	1210	2 A53183	epidermal growth f
9	713.5	23.3	644	2 A36325	epidermal growth f
10	697	22.8	1342	2 A36223	kinase-related tra
11	662	21.6	1339	2 JC4387	epidermal growth f
12	651	21.3	1166	1 S06142	protein-tyrosine k
13	539.5	17.6	843	2 A27131	epidermal growth f
14	404.5	13.2	1323	2 E88257	protein let-23 (im
15	404.5	13.2	1374	2 S70712	protein-tyrosine k
16	390.5	12.8	1369	2 S70713	protein-tyrosine k
17	379	12.4	1330	1 GQFFE	epidermal growth f
18	341	11.1	366	2 D45558	epidermal growth f
19	341	11.1	1717	1 A45558	epidermal growth f
20	331	10.8	333	2 B45558	epidermal growth f
21	331	10.8	342	2 C45558	epidermal growth f
22	318	10.4	1363	2 T42220	insulin-like growt
23	276.5	9.0	1477	2 T18534	protein-tyrosine k
24	271	8.9	2101	2 S57245	insulin receptor (
25	271	8.9	2148	1 A56081	insulin receptor -
26	258	8.4	1382	1 INHUR	insulin receptor p
27	252.5	8.2	1372	2 A34157	insulin receptor p
28	252	8.2	1607	2 T43212	insulin-like growt
29	249	8.1	1383	2 A36080	insulin receptor p

RESULT 1

A24571

protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erb
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth ;
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epider
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491

ALIGNMENTS

30	246.5	8.1	1300	2	A36502	insulin receptor-r
31	238.5	7.8	1367	1	IGHUR1	insulin-like growt
32	236.5	7.7	1371	2	A33837	insulin-like growt
33	235.5	7.7	183	2	JH0803	tyrosine kinase re
34	233.5	7.6	540	2	B47417	insulin receptor-r
35	231.5	7.6	1268	2	B36502	insulin receptor-r
36	224.5	7.3	329	2	A48805	insulin-like growt
37	207	6.8	1390	2	T30346	insulin receptor -
38	188	6.1	1846	2	J42047	insulin receptor h
39	168	5.5	386	1	JH0610	acid phosphatase (
40	155	5.1	5262	2	T03454	ALR protein - huma
41	153.5	5.0	707	2	A46302	PTB-associated spl
42	146	4.8	351	2	S50754	hypothetical prote
43	146	4.8	442	2	S50062	cell wall glycopro
44	139.5	4.6	473	2	S50755	hypothetical prote
45	139.5	4.6	597	2	S72468	probable transcrip

A:Cross-references: GB:L29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A:Cross-references: GB:M16792; NID:q183983; PIDN:AAA58637.1; PID:g553332
C:Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C:Genetics:
A:Gene: GDB:ERBB2; NGL; NEU; HER-2
A:Cross-references: GDB:120613; OMIM:164870
A:Map position: 17q21.1-17q21.1
A:Introns: 25/1; 75/3; 147/1; 883/3
A:Note: the list of introns is incomplete
C:Function:
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
A:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F:22-653/Domain: extracellular #status predicted <EXT>
F:70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F:395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F:654-675/Domain: transmembrane #status predicted <TM>
F:676-1255/Domain: intracellular #status predicted <INT>
F:718-983/Domain: intracellular kinase homology <KIN>
F:726-734/Region: protein kinase ATP-binding motif
F:68.124.187.259.530.571.629/Binding site: carboxylate (Asn) (covalent) #status predicted
F:686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F:753/Active site: Lys #status predicted
F:1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 78.9%; Score 2417; DB 1; Length 1255;
Best Local Similarity 41.2%; Pred. No. 2.5e-135;
Matches 511; Conservative 4; Mismatches 6; Indels 718; Gaps 2;

QY	35	LARGASTQVCTGDKMLRASPETHLDMRLHLYOGQVQGNLELYLPTNASLSFLQ	94
DB	16	LPFGAATQVCTGDKMLRASPETHLDMRLHLYOGQVQGNLELYLPTNASLSFLQ	75
QY	95	DIQEVQGVYLIAHNVROVPLQRLRVGTQLFEDNYALAVLDNGDPLNTPPTVTCASPG	154
DB	76	DIQEVQGVYLIAHNVROVPLQRLRVGTQLFEDNYALAVLDNGDPLNTPPTVTCASPG	135
QY	155	GLRELQLSLTILKGGVLIQRNPOLCYQDTILMKDIFHNKQLALTIDITNRSRACHPC	214
DB	136	GLRELQLSLTILKGGVLIQRNPOLCYQDTILMKDIFHNKQLALTIDITNRSRACHPC	195
QY	215	SPMCKGSRGWSSSDCSLTTCVAGCARGKGLPTDCHEQCAAGCTGPKHSDCLAC	274
DB	196	SPMCKGSRGWSSSDCSLTTCVAGCARGKGLPTDCHEQCAAGCTGPKHSDCLAC	255
QY	275	LHFNHSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSASII-	333
DB	256	LHFNHSGICELHCPALVYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGSCTLVC	315
QY	334	-----	333
DB	316	PLHNQEVTAEDGTORCEKSKPCARVCYGLGHLREVRVTSANIQFAGCKIFGSLA	375
QY	334	-----	333
DB	376	FLPESFDGDPASNTAPLOEQVFPETLEETITGYLYISAWPDLPLSVFQNLQVIRGRI	435
QY	334	-----	333
DB	436	LHNGAYSULTQGLGISWGLRLSRLGSLALIHNTHLCLFVHTVPWDLFPNPHQALLH	495
QY	334	-----	333
DB	496	TANRPEDECVGBGLACHQLCARGHCWGPQTQVNCSTQFRLQGEVCEECVQLGLPREYV	555
QY	334	-----	333
DB	556	NARHCLPCHPECPQNGSVTCFPGPADQCACAHYKDPFCVACPSGVKPDLSYMPIWK	615
QY	334	-----	333

DB	616	FPDEBGAQPCPINCSTHSCVDLDDKGCPAEQRASPLTSIIISAVVGILLVVLGVVFGILI	675
QY	334	-----	333
DB	676	KRQOKIRKYTWRRLLQETELVEPLTPSGAMPNQAMRILKETELURKVVLGSGAFGVY	735
QY	334	-----	333
DB	736	KGIWIPDGENVKI PVAIKVLRNTSPKANKELDEAYVMAGVGSPPVSRLLGICLTSTVQ	795
QY	334	-----	333
DB	796	LVTQLMPYCLLDHVRENRGLRGSQDLLNWCQIAKGMSYLEDVRLVHRDLAARNVLVKS	855
QY	334	-----	333
DB	856	PNHVKITDFGLARLLDIDETEHADGGKVPIKMMALESILRRRFTHQSDVWSYGVTVWEL	915
QY	334	-----NFEKL---	338
DB	916	MTFGAKPYDGI PAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRFRELVSE	975
QY	339	-----	338
DB	976	FARMARDPQRFVVIQNEGLGPASPLDSTFYRSLLEDDMDGLVDAAEYLVFQQGFPCDP	1035
QY	339	---GAGGMVHFRSSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA	396
DB	1036	APGAGMVHFRSSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA	1095
QY	397	KGLQSLPHTDPSLQRYSEDTVPPLPSETDGVVAPLTCSPQPEYVQVDPVRPQPSREG	456
DB	1096	KGLQSLPHTDPSLQRYSEDTVPPLPSETDGVVAPLTCSPQPEYVQVDPVRPQPSREG	1155
QY	457	PLPAARPAAGATLERAKTLSPGKNGVVDVAFAGGAVENPEYLTPOGGAAPQHPPPAFSP	516
DB	1156	PLPAARPAAGATLERAKTLSPGKNGVVDVAFAGGAVENPEYLTPOGGAAPQHPPPAFSP	1215
QY	517	AFDNLVYWDQPPPERGAPSTFKGTPTAENPEYLGLDVP	555
DB	1216	AFDNLVYWDQPPPERGAPSTFKGTPTAENPEYLGLDVP	1254

RESULT 2
TYRWTU
protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C:Accession: A24562; A61204
R:Batgmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A:Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A:Reference number: A24562; MUID:86118662; PMID:3945311
A:Accession: A24562
A:Molecule type: mRNA
A:Residues: 1-1260 <BAR>
A:Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
R:Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
Carcinogenesis 12, 1975-1978, 1991
A:Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no
2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A:Reference number: A61204; MUID:92035293; PMID:1682063
A:Accession: A61204
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 637-663, 'V', 665-702 <MAS>
A:Note: authors translated the codon GCA for residue 25 as Val
C:Genetics:
A:Gene: neu
C:Superfamily: epidermal growth factor receptor; protein kinase homology
C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 59.6409 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-3

Perfect score: 3062

Sequence: 1 MRAAPLLARASLSGLFLF.....ENPEYGLDVPAAHHHHH 564

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3062	100.0	564	US-09-821-883-3	Sequence 3, Appli
2	3014	98.4	697	US-09-821-883-4	Sequence 4, Appli
3	3005.5	98.2	555	US-09-821-883-1	Sequence 1, Appli
4	2957.5	96.6	690	US-09-821-883-2	Sequence 2, Appli
5	2585	84.4	919	US-09-854-356-6	Sequence 6, Appli
6	2422	79.1	1255	US-09-811-123-9	Sequence 9, Appli
7	2422	79.1	1255	US-09-811-115-3	Sequence 3, Appli
8	2422	79.1	1255	US-09-769-508-2	Sequence 2, Appli
9	2422	79.1	1255	US-09-984-092-4	Sequence 4, Appli
10	2422	79.1	1255	US-10-177-293-126	Sequence 126, App
11	2422	79.1	1255	US-10-207-498-6	Sequence 6, Appli
12	2422	79.1	1255	US-10-338-730-2	Sequence 2, Appli
13	2422	79.1	1255	US-10-322-892-4	Sequence 4, Appli
14	2422	79.1	1255	US-10-272-437A-28	Sequence 28, Appl
15	2422	79.1	1255	US-10-117-937-594	Sequence 594, App

ALIGNMENTS

RESULT 1

US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct
US-09-821-883-3

Query Match 100.0%; Score 3062; DB 9; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.3e-206;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRAAPLLARASLSGLFLFLFFWLDRLSVLAKELARGAASVQVCTGTDKMLRLPASPT 60
Db	1	MRAAPLLARASLSGLFLFLFFWLDRLSVLAKELARGAASVQVCTGTDKMLRLPASPT 60
Qy	61	HLDMRLHYQGVQVQGNLELYLPTNASLSFLQDIQVQGVYLAHNOVQVPLQRLRI 120
Db	61	HLDMRLHYQGVQVQGNLELYLPTNASLSFLQDIQVQGVYLAHNOVQVPLQRLRI 120
Qy	121	VRGTQLFEDNVALVDNGDPLNNTPTVTGASPGGLRELQLRSLTEILKGGVLIQRNPOL 180

16	2422	79.1	1255	15	US-10-435-696-36	Sequence 36, Appli
17	2422	79.1	1255	16	US-10-441-779C-4	Sequence 4, Appli
18	2422	79.1	1255	16	US-10-734-564-126	Sequence 126, App
19	2417	78.9	1255	9	US-09-854-356-1	Sequence 1, Appli
20	2417	78.9	1255	9	US-09-930-125-2	Sequence 2, Appli
21	2417	78.9	1255	10	US-09-441-411-6	Sequence 6, Appli
22	2417	78.9	1255	12	US-10-469-162-3	Sequence 3, Appli
23	2417	78.9	1255	12	US-10-253-286-553	Sequence 553, App
24	2417	78.9	1255	12	US-09-765-973-2	Sequence 2, Appli
25	2417	78.9	1255	12	US-10-418-027-3	Sequence 3, Appli
26	2417	78.9	1255	14	US-10-207-655-45	Sequence 45, Appli
27	2417	78.9	1255	14	US-10-313-644-2	Sequence 2, Appli
28	2417	78.9	1255	15	US-10-394-322A-17	Sequence 17, Appli
29	2417	78.9	1255	15	US-10-245-871-553	Sequence 553, App
30	2417	78.9	1255	15	US-10-149-138-4641	Sequence 4641, App
31	2417	78.9	1255	16	US-10-647-005-68	Sequence 68, Appli
32	2417	78.9	1255	16	US-10-149-138-4641	Sequence 4641, App
33	2410	78.7	1253	14	US-10-146-473-72	Sequence 72, Appli
34	1974.5	64.5	1256	9	US-09-854-356-14	Sequence 14, Appli
35	1962	64.1	1260	9	US-09-870-759-118	Sequence 118, App
36	1962	64.1	1260	10	US-09-751-708A-118	Sequence 118, App
37	1959.5	64.0	1256	9	US-09-854-356-2	Sequence 2, Appli
38	1838.5	60.0	479	9	US-09-821-883-5	Sequence 5, Appli
39	1614.5	52.7	645	9	US-09-921-161-1	Sequence 1, Appli
40	1614.5	52.7	645	14	US-10-268-501-13	Sequence 13, Appli
41	1614.5	52.7	645	15	US-10-608-626-13	Sequence 13, Appli
42	1614.5	52.7	653	9	US-09-854-356-3	Sequence 3, Appli
43	1614.5	52.7	685	15	US-10-412-804A-4	Sequence 4, Appli
44	1614.5	52.7	690	15	US-10-412-804A-11	Sequence 11, Appli
45	1614.5	52.7	712	9	US-09-854-356-7	Sequence 7, Appli

Db 121 VRTQTFEDNYALVLDNGDPLNNTTPTVTGASFGGLRELQLRSLTEILKGGVLIQRNPOL 180
Qy 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Qy 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Db 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Qy 361 TLGLEPSEEEAPRSPPLAPSEAGSDVFDGDLGMGAAGKGLQSLPTHDPSPLOQRYSDPTVP 420
Db 361 TLGLEPSEEEAPRSPPLAPSEAGSDVFDGDLGMGAAGKGLQSLPTHDPSPLOQRYSDPTVP 420
Qy 421 LPSETDGVVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Qy 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYYWDQDPPERGAPSTPKG 540
Db 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYYWDQDPPERGAPSTPKG 540
Qy 541 TPTAENPEYLGIDVPAHHHHH 564
Db 541 TPTAENPEYLGIDVPAHHHHH 564

RESULT 2
US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

Query Match 98.4%; Score 3014; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 4.1e-203;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGTMKRLPASPET 60
Db 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGTMKRLPASPET 60
Qy 61 HLDMLRHLVQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIQHNQVQVPLQRLRI 120
Db 61 HLDMLRHLVQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIQHNQVQVPLQRLRI 120
Qy 121 VRTQTFEDNYALVLDNGDPLNNTTPTVTGASFGGLRELQLRSLTEILKGGVLIQRNPOL 180
Db 121 VRTQTFEDNYALVLDNGDPLNNTTPTVTGASFGGLRELQLRSLTEILKGGVLIQRNPOL 180

Qy 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTFES 300
Qy 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Db 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGGMVHRRSSSTRSGGDL 360
Qy 361 TLGLEPSEEEAPRSPPLAPSEAGSDVFDGDLGMGAAGKGLQSLPTHDPSPLOQRYSDPTVP 420
Db 361 TLGLEPSEEEAPRSPPLAPSEAGSDVFDGDLGMGAAGKGLQSLPTHDPSPLOQRYSDPTVP 420
Qy 421 LPSETDGVVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Db 421 LPSETDGVVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAGATLERAKTILSPGKNG 480
Qy 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYYWDQDPPERGAPSTPKG 540
Db 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLYYWDQDPPERGAPSTPKG 540
Qy 541 TPTAENPEYLGIDVPAHHH 558
Db 541 TPTAENPEYLGIDVPAHHH 558

RESULT 3
US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 98.2%; Score 3005.5; DB 9; Length 555;
Best Local Similarity 98.4%; Pred. No. 1.2e-202;
Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;
Qy 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGTMKRLPASPET 60
Db 1 MRAAPLLIARAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGTMKRLPASPET 60
Qy 61 HLDMLRHLVQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIQHNQVQVPLQRLRI 120
Db 61 HLDMLRHLVQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIQHNQVQVPLQRLRI 120
Qy 121 VRTQTFEDNYALVLDNGDPLNNTTPTVTGASFGGLRELQLRSLTEILKGGVLIQRNPOL 180
Db 121 VRTQTFEDNYALVLDNGDPLNNTTPTVTGASFGGLRELQLRSLTEILKGGVLIQRNPOL 180
Qy 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNQALALTLIDNRSRACHPCSPCKGRCWGSESSDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 16.4294 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYLGLDVPAHHHHH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2422	79.1	1255	4	US-09-527-487-2
2	2422	79.1	1255	4	US-09-811-115-3
3	2417	78.9	1255	1	US-08-467-083-68
4	2417	78.9	1255	1	US-08-414-417B-68
5	2417	78.9	1255	2	US-08-486-348A-68
6	2417	78.9	1255	2	US-08-625-101-2
7	2417	78.9	1255	2	US-08-468-545B-68
8	2417	78.9	1255	2	US-08-356-786-2
9	2417	78.9	1255	3	US-08-466-680B-68
10	2417	78.9	1255	4	US-09-354-533-68
11	2397	78.3	1255	2	US-08-484-438-8
12	1614.5	52.7	782	2	US-09-146-283-4
13	1614.5	52.7	782	3	US-08-579-823A-4
14	1614.5	52.6	419	4	US-09-630-195-4
15	1610.5	52.6	624	3	US-09-344-155-2
16	1599.5	52.2	624	3	US-08-422-108-1
17	1599.5	52.2	624	4	US-08-422-734-1
18	1177	38.4	580	1	US-08-414-417B-69
19	1177	38.4	580	2	US-08-486-348A-69
20	1177	38.4	580	2	US-08-468-545B-69
21	1177	38.4	580	3	US-08-466-680B-69
22	1177	38.4	580	4	US-09-354-533-69
23	860	28.1	166	4	US-09-648-067A-1
24	723.5	23.6	911	2	US-08-484-438-10
25	723.5	23.6	1058	2	US-08-484-438-4
26	723.5	23.6	1308	2	US-08-484-438-2
27	719.5	23.5	644	1	US-08-336-708A-9

28	719.5	23.5	1210	2	US-08-484-438-7	Sequence 7, Appli
29	719.5	23.5	1210	2	US-08-475-035-4	Sequence 4, Appli
30	701.5	22.9	478	4	US-09-570-454-2	Sequence 2, Appli
31	701.5	22.9	478	4	US-09-867-521-2	Sequence 2, Appli
32	697	22.8	1342	1	US-07-978-895-4	Sequence 4, Appli
33	697	22.8	1342	2	US-08-484-438-9	Sequence 9, Appli
34	697	22.8	1342	2	US-08-473-119-4	Sequence 4, Appli
35	697	22.8	1342	2	US-08-473-352-4	Sequence 4, Appli
36	697	22.8	1342	4	US-09-170-699-4	Sequence 4, Appli
37	697	22.8	1343	6	5183884-4	Patent No. 5183884
38	493	16.1	97	1	US-08-421-356-3	Sequence 3, Appli
39	493	16.1	97	4	US-09-046-783-3	Sequence 3, Appli
40	259	8.5	1382	2	US-08-737-715-2	Sequence 2, Appli
41	259	8.5	1382	2	US-09-457-040B-7	Sequence 7, Appli
42	238.5	7.8	516	3	US-08-746-559A-4	Sequence 4, Appli
43	238.5	7.8	1367	2	US-08-249-687C-2	Sequence 2, Appli
44	238.5	7.8	1367	2	US-08-625-819-2	Sequence 2, Appli
45	238.5	7.8	1367	3	US-08-746-559A-2	Sequence 2, Appli

RESULT 1
US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match 79.1%; Score 2422; DB 4; Length 1255;
Best Local Similarity 41.3%; Pred. No. 1.3e-187;
Matches 512; Conservative 4; Mismatches 5; Indels 718; Gaps 2;

ALIGNMENTS

```
QY 334 ----- 333
Db 436 LHNGAYSLTLOGLGISWGLRSLRSLGSLALIHNNHNLCPVHTVPWDLFRNPHQALLH 495
QY 334 ----- 333
Db 496 TANRPEDECVEGLACHQLCARGHCWGPPTQCVNCSQFLRQECVEBECVRLQGLPREVY 555
QY 334 ----- 333
Db 556 NARHCLPCHPECPONGSVTCFGEADQCVACAHYKDPFCVRCPSGVKPDLSYMPIWK 615
QY 334 ----- 333
Db 616 FPDEGACQPCPINCTHSCVDLDDKGCAPAEQASPLTSIVSAVVGILLVVVLGVVFGILI 675
QY 334 ----- 333
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQRILKETELRKVKVLGSGAGFTVY 735
QY 334 ----- 333
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKELDEAYVMAGVGSPPVSRLLIGLICLTSTVQ 795
QY 334 ----- 333
Db 796 LVTQLMPYGCLLDHYENRGRGLSQDLLNWCQIAKMSYLEDVLRVHRDLAARNVLKVS 855
QY 334 ----- 333
Db 856 PNHVKITDGLARLLDIDETEHADGKGVPIKWMMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPPICTIDVYIMVKWMIDSECRPRPRELVSE 975
QY 339 ----- 338
Db 976 FSRMARQRFVVIQNEIDLGPASPLDSTFYRSLLEDDMDGLVDAAEYLVQQGFFCPDP 1035
QY 339 --GAGGMVHRRHSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA 396
Db 1036 APGAGMVHRRHSSTSGGDLTLGLEPSEEAAPRSLAPSEGAGSDVFDGLGMGAA 1095
QY 397 KGLQSLPHDPSPLQRYSEDPTVLPSETDGVVAPLTCSPQPEYVQPDVVRPQPSREG 456
Db 1096 KGLQSLPHDPSPLQRYSEDPTVLPSETDGVVAPLTCSPQPEYVQPDVVRPQPSREG 1155
QY 457 PLPAARPAATLERAKTLSPGKNGVVKVDFAFGGAVENPEYLTPQGGAAPOPHPPAFSP 516
Db 1156 PLPAARPAATLERAKTLSPGKNGVVKVDFAFGGAVENPEYLTPQGGAAPOPHPPAFSP 1215
QY 517 AFDNLYYWDQPPPERGAPSTFKGPTTAENPEYLGLDVP 555
Db 1216 AFDNLYYWDQPPPERGAPSTFKGPTTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT. 034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match      79.1%; Score 2422; DB 4; Length 1255;
Best Local Similarity 41.3%; Pred. No. 1.3e-187;
Matches 512; Conservative 4; Mismatches 5; Indels 718; Gaps 2;

QY 35 LARGAASQTQCTGDMKRLRLPASBETHLDMRLHLYQGCQVQGNLELYLPTNLSLSLQ 94
Db 16 LPPGAASQTQCTGDMKRLRLPASBETHLDMRLHLYQGCQVQGNLELYLPTNLSLSLQ 75
QY 95 DIQEVQGVVLAHNOVQVPLQRLRIVRGTQLFEDNYALAVLDNGDPLNNTTPVTGASPG 154
Db 76 DIQEVQGVVLAHNOVQVPLQRLRIVRGTQLFEDNYALAVLDNGDPLNNTTPVTGASPG 135
QY 155 GLRELQRLSRTEILKGGVLIQORNPQLCYQDITLWKDIFHKNNQLALTLIDTNRSRACHPC 214
Db 136 GLRELQRLSRTEILKGGVLIQORNPQLCYQDITLWKDIFHKNNQLALTLIDTNRSRACHPC 195
QY 215 SPMCKGRCWGESSEDCQSLTRTVACGACRCKGPLPTDCCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGRCWGESSEDCQSLTRTVACGACRCKGPLPTDCCHEQCAAGCTGPKHSDCLAC 255
QY 275 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSASII- 333
Db 256 LHFNHSGICELHCPALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
QY 334 ----- 333
Db 316 PLHNQEVTAEDGTQRCCKSKPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLA 375
QY 334 ----- 333
Db 376 FLPSFDDGPASNTAPLQPEQLQVFETLEBITGYLIYISAWPDSLPLDSVFQNLQVIRGI 435
QY 334 ----- 333
Db 436 LHNGAYSLTLOGLGISWGLRSLRSLGSLALIHNNHNLCPVHTVPWDLFRNPHQALLH 495
QY 334 ----- 333
Db 496 TANRPEDECVEGLACHQLCARGHCWGPPTQCVNCSQFLRQECVEBECVRLQGLPREVY 555
QY 334 ----- 333
Db 556 NARHCLPCHPECPONGSVTCFGEADQCVACAHYKDPFCVRCPSGVKPDLSYMPIWK 615
QY 334 ----- 333
Db 616 FPDEGACQPCPINCTHSCVDLDDKGCAPAEQASPLTSIVSAVVGILLVVVLGVVFGILI 675
QY 334 ----- 333
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAQRILKETELRKVKVLGSGAGFTVY 735
QY 334 ----- 333
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKELDEAYVMAGVGSPPVSRLLIGLICLTSTVQ 795
QY 334 ----- 333
Db 796 LVTQLMPYGCLLDHYENRGRGLSQDLLNWCQIAKMSYLEDVLRVHRDLAARNVLKVS 855
QY 334 ----- 333
Db 856 PNHVKITDGLARLLDIDETEHADGKGVPIKWMMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 334 ----- 338
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPPICTIDVYIMVKWMIDSECRPRPRELVSE 975
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 62.1165 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-3
Perfect score: 3062
Sequence: 1 MRAAPLLARASLSGLF.....ENPEYGLDVPAAAHHRHH 564

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107.

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003s.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3062	100.0	564	AAE13110	Human HER
2	3014	98.4	697	AAE13111	Human HER
3	3005.5	98.2	555	AAE13108	Human HER
4	2957.5	96.6	690	AAE13109	Human HER
5	2585	84.4	919	AAE21203	Human HER
6	2585	84.4	919	AAE51148	Her-2/neu
7	2422	79.1	1255	AAE92620	Human her
8	2422	79.1	1255	AAE60167	HER2 tran
9	2422	79.1	1255	AAE12130	Human tyr
10	2422	79.1	1255	AAE26349	Human HER
11	2422	78.1	1255	AAE26366	Human HER
12	2422	79.1	1255	AAU74545	Human HER
13	2422	79.1	1255	ABR47447	Breast ca
14	2422	79.1	1255	ABP74708	Human HER
15	2422	79.1	1255	AAE38390	Human c-e
16	2422	79.1	1255	ADA38143	Human erb
17	2422	79.1	1255	ADA37255	Human Erb
18	2422	79.1	1255	ADB67621	Human epi
19	2417	78.9	1255	AAW01111	HER-2/neu
20	2417	78.9	1255	AAW92406	Human HER
21	2417	78.9	1255	AAE4780	Amino aci
22	2417	78.9	1255	AAE21198	Human HER
23	2417	78.9	1255	AAE88267	HER2/neu
24	2417	78.9	1255	AAE85458	Human HER
25	2417	78.9	1255	AAE20479	Human HER

26	2417	78.9	1255	5	AAU77114	Human Her
27	2417	78.9	1255	5	AAE51143	Human Her
28	2417	78.9	1255	5	AAE24067	Human Her
29	2417	78.9	1255	6	ABR43687	Human c-e
30	2417	78.9	1255	7	ABR82066	Human Her
31	2417	78.9	1255	7	ADC09593	Her2/Neu
32	2417	78.9	1255	7	ADD25484	Binding d
33	2417	78.9	1255	7	ADE63281	Human Pro
34	2417	78.9	1255	7	ADE76190	Human HER
35	2410	78.7	1253	7	ADC35106	Human bre
36	2401	78.4	1433	2	AAE39568	Sequence
37	2346	76.6	1223	5	AAU98923	Human bre
38	2142.5	70.0	920	5	AAE51152	Mouse Her
39	2142.5	70.0	926	5	AAE51153	Mouse Her
40	2104	68.7	1200	3	AAE21208	Human HER
41	1974.5	64.5	1256	3	AAE21206	Mouse Her
42	1974.5	64.5	1256	4	AAE62860	Amino aci
43	1974.5	64.5	1256	5	AAE51151	Mouse Her
44	1962	64.1	1260	6	ABU79135	Erbb2 (HE
45	1959.5	64.0	1256	3	AAE21199	Rat HER-2

ALIGNMENTS

RESULT 1
AAE13110
ID AAE13110 standard; protein; 564 AA.
XX
AC AAE13110;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
XX
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; OVA;
KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
XX
OS Homo sapiens.
OS Unidentified.
OS Chimeric.
XX
PN WO200174855-A2.
XX
PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010515.
XX
XX 30-MAR-2000; 2000US-0193504P.
XX
XX (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddis T;
XX WPI: 2001-662965/76.
XX N-PSDB; AAD21566.

An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

Claim 7; Page 26; 59pp; English.

The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory domain derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are

used to treat cancer e.g. breast carcinoma, ovarian and colon cancer associated with a particularly antigen. The present sequence is HER500 fusion protein construct which comprises human PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal sequence, mature HER-2 membrane distal extracellular domain, an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal intracellular domain and a C-terminal tag

Sequence 564 AA;

Query Match 100.0%; Score 3062; DB 4; Length 564;
Best Local Similarity 100.0%; Pred. No. 1.8e-211;
Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGMDKMLRLPASPET 60
DB 1 MRAAPLLARAAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGMDKMLRLPASPET 60

QY 61 HLDMLRHLYQGCQVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQLRLI 120
DB 61 HLDMLRHLYQGCQVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQLRLI 120

QY 121 VRGTLFEDNALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRPQL 180
DB 121 VRGTLFEDNALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRPQL 180

QY 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240
DB 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240

QY 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
DB 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300

QY 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGMWHRRSSSTRSGGDL 360
DB 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGMWHRRSSSTRSGGDL 360

QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSIPTHDPSPLOYSDDPTVP 420
DB 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSIPTHDPSPLOYSDDPTVP 420

QY 421 LPSETDGYVAPLTCSPQPEYVQNPQVPRPQPSREGPLPAARPAGATLERAKTLSPGKNG 480
DB 421 LPSETDGYVAPLTCSPQPEYVQNPQVPRPQPSREGPLPAARPAGATLERAKTLSPGKNG 480

QY 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPPSTPKG 540
DB 481 VVKOVFAFGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPERGAPPSTPKG 540

QY 541 TPTAENPEYLGIDVPAHAHHHHH 564
DB 541 TPTAENPEYLGIDVPAHAHHHHH 564

RESULT 2

AAE13111
ID AAE13111 standard; protein; 697 AA.
XX AAE13111;
AC AAE13111;
XX 28-JAN-2002 (first entry)
XX Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.
DE Immunostimulatory fusion protein; IFP; antigen component; therapy;
XX immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.

OS Homo sapiens.
OS Rattus norvegicus.
OS Unidentified.
OS Chimeric.
XX WO200174855-A2.
PN 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US010515.
PF 30-MAR-2000; 2000US-0193504P.
PR (DEND-) DENDREON CORP.
PA Laus R, Vidovic D, Graddis T;
PI WPI; 2001-662965/76.
XX N-PSDB; AAD21567.
DR An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.
PT Claim 7; Page 27; 59pp; English.
PS The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are used to treat cancer e.g. breast carcinoma, ovarian and colon cancer associated with a particularly antigen. The present sequence is HER500 rGM-CSF fusion protein construct which comprises human PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal sequence, mature HER-2 membrane distal extracellular domain, an Ala linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal intracellular domain, an Ala linker, a mature rat granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a C-terminal tag

Sequence 697 AA;

Query Match 98.4%; Score 3014; DB 4; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.7e-208;
Matches 558; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARAAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGMDKMLRLPASPET 60
DB 1 MRAAPLLARAAASLSGLFLLFFWLDRLSVLAKELARGAASQVCTGMDKMLRLPASPET 60

QY 61 HLDMLRHLYQGCQVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQLRLI 120
DB 61 HLDMLRHLYQGCQVQVQGNLELTPTNASLSFLQDIQEVQGVLIHNRQVPLQLRLI 120

QY 121 VRGTLFEDNALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRPQL 180
DB 121 VRGTLFEDNALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRPQL 180

QY 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240
DB 181 CYQDTILWKDIFHKNQNLALTIDNRSRACHPCSPMCKSGRCWGESSEDCQSLTRTVCA 240

QY 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
DB 241 GGCARCKGPLETDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300

QY 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGMWHRRSSSTRSGGDL 360
DB 301 MPNPEGRYTFGASCVTACPYNYLSTDVGSASIIINFEKLGAGMWHRRSSSTRSGGDL 360

QY 361 TLGLEPSEEEAPRSLAPSEAGSDVFDGLGMGAAGKLSIPTHDPSPLOYSDDPTVP 420

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 4652.68 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPPAHHHHH 690
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Command line parameters: -MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09821883/runat_09092004_105127_6643/app_query.fasta_1.3100
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEXT=0 -TRANS=human40.cdi -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRN_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -THR SCORE=500 -THR MAX=100 -THR MIN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1 1 11571 @runat_09092004_105127_6643 -NCPU=6 -ICPU=3
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estini.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hci.*
9: gb_est1.*
10: gb_est2.*
11: gb_hci.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_ges_hum.*
18: em_ges_inv.*
19: em_ges_pln.*
20: em_ges_vrt.*
21: em_ges_fun.*
22: em_ges_mam.*
23: em_ges_mus.*
24: em_ges_pro.*
25: em_ges_rtd.*
26: em_ges_phg.*
27: em_ges_vrl.*
28: gb_ges1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1961	52.4	4323	11	AK031099 Mus muscu
2	1944.5	52.0	4463	11	AK083669 Mus muscu
3	1288	34.4	1016	12	BM562913 AGNCOURT
4	1176.5	31.5	1129	12	BM802792 AGNCOURT
5	1175.5	31.4	885	14	CA455074 AGNCOURT
6	1170	31.3	871	14	CA488274 AGNCOURT
7	1163	31.1	4715	11	AF318349 Homo sapi
8	1145.5	30.6	893	14	CA455141 AGNCOURT
9	1098.5	29.4	894	14	CA454131 AGNCOURT
10	1096	29.3	1004	14	CD515356 AGNCOURT
11	1087	29.1	653	12	BM721340 UI-E-E01-
12	1079	28.9	902	14	CA488868 AGNCOURT
13	1073	28.7	588	13	EX478931 DKP2p686A
14	1046	28.0	569	9	AL701765 DKP2p686P
15	1027	27.5	583	13	EX479114 DKP2p686P
16	1016.5	27.2	795	14	CA328613 UI-M-PY0-
17	1009.5	27.0	3110	11	AK031542 Mus muscu
18	1004	26.9	808	14	CA489534 AGNCOURT
19	997	26.7	616	14	CD721801 OJ01f07.Y
20	976	26.1	943	13	BQ958632 AGNCOURT
21	974	26.0	1988	11	BC023725 Mus muscu
22	974	26.0	3372	11	BC046553 Mus muscu
23	933	25.0	855	13	BU594980 AGNCOURT
24	924	24.7	813	12	BI557797 603236977
25	915.5	24.5	852	14	CD516283 AGNCOURT
26	915	24.5	614	10	AW370693 QVI-BT026
27	912	24.4	791	12	BI154872 602902857
28	903	24.2	717	14	CB598701 AGNCOURT
29	901	24.1	491	12	BM790293 K-EST0070
30	893	23.9	906	14	CA454570 AGNCOURT
31	889.5	23.8	998	12	BI649877 603296516
32	881	23.6	685	9	AU123871 AU123871
33	874.5	23.4	932	14	CA487981 AGNCOURT
34	868	23.2	609	10	AW701942 uq93b02.Y
35	864.5	23.1	621	12	BG283493 602407782
36	853	22.8	887	13	BQ717097 AGNCOURT
37	848	22.7	786	12	BI155788 602904360
38	835.5	22.3	649	9	AA496412 2v37a02.r
39	817	21.9	533	14	CB437620 685491.MA
40	817	21.9	552	12	BG277542 ux45d07.Y
41	811	21.7	670	9	AI906012 RC-BT105-
42	808	21.6	876	13	BQ769889 UI-M-PT0-
43	803	21.5	887	14	CA980253 AGNCOURT
44	802	21.4	514	10	BF998814 QVO-GN014
45	792.5	21.2	678	12	BI555157 603236396

ALIGNMENTS

RESULT 1
AK031099
LOCUS
DEFINITION Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:530404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.
ACCESSION AK031099
VERSION AK031099.1 GI:26082143
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AK031099 4323 bp mRNA linear HTC 18-SEP-2003

```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yanamoto, R., Maeumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
FEATURES
source
/db_xref="FANTOM DB:5930404N10"
/db_xref="MGI:2394277"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
misc_feature
48. _3818
/notes="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2,
neuro/glioblastoma derived oncogene homolog (avian)
(MGI:MGI:95410, GB|U71126, evidence: BLASTN, 99%,
match=449)"
ORIGIN
Alignment Scores:
Pred. No.: 3,81e-107 Length: 4323
Score: 1961.00 Matches: 455
Percent Similarity: 36.01% Conservative: 33
Best Local Similarity: 33.58% Mismatches: 103
Query Match: 52.45% Indels: 764
DB: 11 Gaps: 6
US-09-821-883-2 (1-690) x AK031099 (1-4323)
Qy 3 AlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeuLeu 22
Db 33 GCCGAGTATCATCATGGAGCTGGCGCTGGTGGGGGTTCCTCTCGCCCTC 92
Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
Db 93 -----CTGTCCCCCGGAGCGCGGTACC 116
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
Db 117 CAAGTGTGTACCGGTACCGACATGAGTTGGACTCCCTGCCAGTCTCTGAGACCCACTG 176
Qy 63 AspMetLeuArgHisLeuTyrglnGlyCysGlnValValGlnGlnGlyAsnLeuGluLeuThr 82
Db 177 GACATGTTGCCACCTCTACAGGGCTGTGAGTGTGTGAGGGCAATTTGGAGCTTACC 236
Qy 83 TyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGlyTyr 102
Db 237 TACCTGCCCGCAATGCCAGCTCTCATCTCTGCGAGGACATCCAGGAAGTCCAGGGATAC 296
Qy 103 ValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIleValArg 122
Db 297 ATGCTCATCGTCTCAACACCGAGTGAAACACGTCCTCCACTGCAGAGGTTGCGCATCGTGAGA 356
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAspProLeu 142
Db 357 GGGACTCAGCTCTTTGAGGACAGTATGCCCTGGCTGTGTAGACACCGAGACCTTTG 416
Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGlnLeu 161
Db 417 GACAACTGTCACCGCGCCCGCCAGGAGAGAACCCAGAGGGCTGCGGAGCTGCAGCTT 476
Qy 162 ArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeuCys 181
Db 477 CGAAGTCTCAGAGATCTTGAAGGGAGGAGTTTGTATCGGTGGGAACCCCTCAGCTCTGC 536
Qy 182 TyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
Db 537 TACGAGACATGGTTTGTGGAGGATGTCTCCGTAAAGATATACAGCTGGCTCCTCTC 596
Qy 202 LeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
Db 597 GACATGGACCAATCGTCCCGGCTGTCCACTGTGTGTGCCCCCAACCTGCAAGACAAT 656
Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAlaGly 241
Db 657 CACTGTTGGGGTGAGAGTCTCTGAAGACTGTGAGATCTTGTGACCATCTGTACTAGT 716

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 834.278 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLIARAASLSGLF.....EPVOEGAPPPPAHHHHH 690

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0			0.5
Fgapop 6.0			7.0
Delop 6.0			7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09821883/runat_09092004_105129_6736/app_query.fasta_1.3100
-DB=Published Applications NA -QFMT=fastap -SUFFX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=ptc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09821883 @CGN 1 1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/cgn2_6/ptodata/1/pubna/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/1/pubna/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/1/pubna/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/1/pubna/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/1/pubna/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/1/pubna/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/1/pubna/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq2.*
14:	/cgn2_6/ptodata/1/pubna/US10A_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/1/pubna/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/1/pubna/US10C_PUBCOMB.seq.*
17:	/cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq2.*
19:	/cgn2_6/ptodata/1/pubna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

ALIGNMENTS

RESULT 1

```
US-09-821-883-7
; Sequence 7, Application US/09821883
; Patent No. US2002061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2070
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-hGM-CSF construct
```

1	3739	100.0	2070	9	US-09-821-883-7	Sequence 7, Appl
2	3473.5	92.9	2091	9 <td>US-09-821-883-9</td> <td>Sequence 9, Appl</td>	US-09-821-883-9	Sequence 9, Appl
3	2974	79.5	1665	9 <td>US-09-821-883-6</td> <td>Sequence 6, Appl</td>	US-09-821-883-6	Sequence 6, Appl
4	2959.5	79.2	1692	9 <td>US-09-821-883-8</td> <td>Sequence 8, Appl</td>	US-09-821-883-8	Sequence 8, Appl
5	2426	64.9	4543	9 <td>US-09-769-508-1</td> <td>Sequence 1, Appl</td>	US-09-769-508-1	Sequence 1, Appl
6	2423	64.8	4530	9 <td>US-09-877-177-11</td> <td>Sequence 11, Appl</td>	US-09-877-177-11	Sequence 11, Appl
7	2423	64.8	4530	13	US-10-392-113-45	Sequence 45, Appl
8	2423	64.8	4530	13	US-10-426-836-11	Sequence 11, Appl
9	2423	64.8	4530	15	US-10-177-293-125	Sequence 125, App
10	2423	64.8	4530	15	US-10-007-926A-119	Sequence 119, App
11	2423	64.8	4530	15	US-10-338-730-1	Sequence 1, Appl
12	2423	64.8	4530	15	US-10-101-510-124	Sequence 124, App
13	2423	64.8	4530	16	US-10-116-275-131	Sequence 131, App
14	2423	64.8	4530	16	US-10-272-437A-27	Sequence 27, Appl
15	2423	64.8	4530	16	US-10-117-937-595	Sequence 595, App
16	2423	64.8	4530	16	US-10-159-563-208	Sequence 208, App
17	2423	64.8	4530	16	US-10-435-696-10	Sequence 10, Appl
18	2423	64.8	4530	17	US-10-734-564-59	Sequence 59, Appl
19	2423	64.8	4642	15	US-10-198-846-10896	Sequence 10896, A
20	2422	64.8	9274	9 <td>US-09-811-123-7</td> <td>Sequence 7, Appl</td>	US-09-811-123-7	Sequence 7, Appl
21	2422	64.8	9274	9 <td>US-09-811-115-1</td> <td>Sequence 1, Appl</td>	US-09-811-115-1	Sequence 1, Appl
22	2418	64.7	4473	10 <td>US-09-441-411-5</td> <td>Sequence 5, Appl</td>	US-09-441-411-5	Sequence 5, Appl
23	2418	64.7	4473	15 <td>US-10-146-473-32</td> <td>Sequence 32, Appl</td>	US-10-146-473-32	Sequence 32, Appl
24	2418	64.7	4473	15 <td>US-10-207-655-44</td> <td>Sequence 44, Appl</td>	US-10-207-655-44	Sequence 44, Appl
25	2418	64.7	4473	15 <td>US-10-101-510-81</td> <td>Sequence 81, Appl</td>	US-10-101-510-81	Sequence 81, Appl
26	2410.5	64.5	3765	15 <td>US-10-207-498-5</td> <td>Sequence 5, Appl</td>	US-10-207-498-5	Sequence 5, Appl
27	2410.5	64.5	3768	9 <td>US-09-811-123-8</td> <td>Sequence 8, Appl</td>	US-09-811-123-8	Sequence 8, Appl
28	2410.5	64.5	3768	9 <td>US-09-811-115-2</td> <td>Sequence 2, Appl</td>	US-09-811-115-2	Sequence 2, Appl
29	2410.5	64.5	3768	13 <td>US-09-984-092-3</td> <td>Sequence 3, Appl</td>	US-09-984-092-3	Sequence 3, Appl
30	2410.5	64.5	3768	13 <td>US-10-280-576-3</td> <td>Sequence 3, Appl</td>	US-10-280-576-3	Sequence 3, Appl
31	2410.5	64.5	3768	17 <td>US-10-441-793-3</td> <td>Sequence 3, Appl</td>	US-10-441-793-3	Sequence 3, Appl
32	2405.5	64.3	3768	9 <td>US-09-854-356-9</td> <td>Sequence 9, Appl</td>	US-09-854-356-9	Sequence 9, Appl
33	2405.5	64.3	3768	9 <td>US-09-930-125-1</td> <td>Sequence 1, Appl</td>	US-09-930-125-1	Sequence 1, Appl
34	2405.5	64.3	3768	13 <td>US-09-765-973-1</td> <td>Sequence 1, Appl</td>	US-09-765-973-1	Sequence 1, Appl
35	2405.5	64.3	3768	15 <td>US-10-313-644-1</td> <td>Sequence 1, Appl</td>	US-10-313-644-1	Sequence 1, Appl
36	2393	64.0	4606	10 <td>US-09-971-392-70</td> <td>Sequence 70, Appl</td>	US-09-971-392-70	Sequence 70, Appl
37	2188.5	58.5	1437	9 <td>US-09-821-883-10</td> <td>Sequence 10, Appl</td>	US-09-821-883-10	Sequence 10, Appl
38	1968	52.6	3955	9 <td>US-09-870-759-117</td> <td>Sequence 117, App</td>	US-09-870-759-117	Sequence 117, App
39	1968	52.6	3955	9 <td>US-09-854-356-10</td> <td>Sequence 10, Appl</td>	US-09-854-356-10	Sequence 10, Appl
40	1968	52.6	3955	10 <td>US-09-751-708A-117</td> <td>Sequence 117, App</td>	US-09-751-708A-117	Sequence 117, App
41	1961	52.4	3771	9 <td>US-09-854-356-11</td> <td>Sequence 11, Appl</td>	US-09-854-356-11	Sequence 11, Appl
42	1684.5	44.5	2132	16 <td>US-10-412-804A-3</td> <td>Sequence 3, Appl</td>	US-10-412-804A-3	Sequence 3, Appl
43	1659	44.4	2164	16 <td>US-10-412-804A-5</td> <td>Sequence 5, Appl</td>	US-10-412-804A-5	Sequence 5, Appl
44	1656	44.3	2149	16 <td>US-10-412-804A-9</td> <td>Sequence 9, Appl</td>	US-10-412-804A-9	Sequence 9, Appl
45	1587	42.4	867	9 <td>US-09-821-883-24</td> <td>Sequence 24, Appl</td>	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-7

Alignment Scores:

Pred. No.:	0	Length:	2070
Score:	3739.00	Matches:	690
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-821-883-2 (1-690) x US-09-821-883-7 (1-2070)

Qy	1	Met	Arg	Ala	Ala	Pro	Leu	Leu	Leu	Ala	Arg	Ala	Ala	Ser	Leu	Ser	Leu	Gly	Phe	Leu	Phe	20	
Db	1	ATG	AGC	TGC	ACC	CC	TCT	CTC	TGT	GC	CA	GGG	CAG	CA	AGC	CTT	TAG	CCT	TGG	CTT	CTG	TTT	60
Qy	21	Leu	Leu	Phe	Phe	Trp	Leu	Asp	Arg	Ser	Val	Leu	Ala	Leu	Ser	Glu	Leu	Ala	Arg	Gly	Val	Ala	40
Db	61	CTG	CTT	TTTT	TTT	TCT	TG	CT	TAG	AC	CGA	AGT	TAC	TAG	CCA	AGG	AGT	TG	TGC	GGC	GGG	CGG	120
Qy	41	Ser	Thr	Gln	Val	Cys	Thr	Gly	Thr	Asp	Met	Leu	Ser	Leu	Arg	Leu	Pro	Ala	Ser	Pro	Glu	Thr	60
Db	121	TCG	ACC	CCA	AGT	TGC	AC	CGG	CAC	ACAT	GA	AGC	TGA	GGC	TCT	CCG	TCC	AGT	CCC	GAG	ACC	180	
Qy	61	His	Leu	Asp	Met	Leu	Arg	His	Leu	Tyr	Gln	Gly	Cys	Gln	Val	Val	Gln	Gly	Asn	Leu	Glu	80	
Db	181	CAC	CTG	GAC	ATG	CT	CCG	CCA	CC	CTC	TAC	AGG	GCT	GC	AGG	TGT	GC	AGG	AA	AA	CTG	GAA	240
Qy	81	Leu	Thr	Tyr	Leu	Pro	Thr	Asn	Ala	Ser	Leu	Ser	Phe	Leu	Gln	Asp	Ile	Gln	Leu	Val	Gln	100	
Db	241	CTC	ACC	TAC	CTG	CC	CA	CA	ATG	CC	CA	AGC	TGT	CT	TCT	TGC	AGG	AT	TAC	GAG	GTG	CAG	300
Qy	101	Gly	Tyr	Val	Leu	Ile	Ala	His	Asn	Gln	Val	Arg	Gln	Val	Pro	Leu	Gln	Arg	Leu	Arg	Ile	120	
Db	301	GGC	TAC	TGT	CTC	AT	CG	CTC	CA	ACC	AGT	GA	GGC	AGG	TCT	CCC	ACT	GC	AGG	CGT	CGG	ATT	360
Qy	121	Val	Arg	Gly	Thr	Gln	Leu	Phe	Glu	Asp	Asn	Tyr	Ala	Leu	Ala	Val	Leu	Asp	Asn	Gly	Asp	140	
Db	361	GTG	CGA	GGC	CAC	CC	AGC	CT	CTT	TG	AG	GACA	AACT	AT	GCC	TGG	CCG	TG	CTA	GACA	ATG	GAG	420
Qy	141	Pro	Leu	Asn	Asn	Thr	Pro	Val	Thr	Gly	Ala	Ser	Pro	Gly	Gly	Leu	Arg	Glu	Leu	Gln	160		
Db	421	CCG	CTG	AA	CAAT	TAC	CAC	CC	CT	TGT	CA	CG	GGG	CT	CCC	CAG	AGG	CTG	CGG	AGG	TCG	CAG	480
Qy	161	Leu	Arg	Ser	Leu	Thr																	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 118.947 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPAAHHHHH 690

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105128_6657/app_query.fasta_1.3100
-DB=Issued Patents NA -QWTF=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOFCLO=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @C@N 1.1 215 @runat_09092004_105128_6657 -NCPUS=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2423	64.8	4530	1	US-08-229-515A-9
2	2423	64.8	4530	1	US-08-229-515A-9
3	2423	64.8	4530	4	US-08-645-865-9
4	2423	64.8	4530	4	US-09-167-322-4
5	2423	64.8	4530	4	US-09-527-487-1
6	2423	64.8	4530	4	US-09-877-177A-11
7	2418	64.7	4473	4	US-09-811-115-1
8	2418	64.7	4473	3	US-09-048-804-1
9	2418	64.7	4473	3	US-09-056-105-26
10	2410.5	64.5	3768	4	US-09-663-834A-3
11	2405.5	64.3	3768	4	US-09-811-115-2
12	2405.5	64.3	3768	2	US-08-625-101-1
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 4, Appli
					Sequence 5, Appli
					Sequence 6, Appli
					Sequence 7, Appli
					Sequence 8, Appli
					Sequence 9, Appli
					Sequence 10, Appli
					Sequence 11, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 14, Appli
					Sequence 15, Appli
					Sequence 16, Appli
					Sequence 17, Appli
					Sequence 18, Appli
					Sequence 19, Appli
					Sequence 20, Appli
					Sequence 21, Appli
					Sequence 22, Appli
					Sequence 23, Appli
					Sequence 24, Appli
					Sequence 25, Appli
					Sequence 26, Appli
					Sequence 27, Appli
					Sequence 28, Appli
					Sequence 29, Appli
					Sequence 30, Appli
					Sequence 31, Appli
					Sequence 32, Appli
					Sequence 33, Appli
					Sequence 34, Appli
					Sequence 35, Appli
					Sequence 36, Appli
					Sequence 37, Appli
					Sequence 38, Appli
					Sequence 39, Appli
					Sequence 40, Appli
					Sequence 41, Appli
					Sequence 42, Appli
					Sequence 43, Appli
					Sequence 44, Appli
					Sequence 45, Appli

13	2258.5	60.4	2385	2	US-09-146-283-3	Sequence 3, Appli
14	2258.5	60.4	2385	3	US-08-579-823A-3	Sequence 3, Appli
15	2258.5	60.4	2385	3	US-09-344-195-3	Sequence 3, Appli
16	1968	52.6	3955	1	US-08-229-515A-14	Sequence 14, Appli
17	1968	52.6	3955	1	US-08-645-865-14	Sequence 14, Appli
18	1647	44.0	1872	3	US-08-422-108-2	Sequence 2, Appli
19	1647	44.0	1872	4	US-08-422-734-2	Sequence 2, Appli
20	770	20.6	1588	2	US-09-146-283-1	Sequence 1, Appli
21	770	20.6	1588	3	US-08-579-823A-1	Sequence 1, Appli
22	770	20.6	1588	3	US-09-344-195-1	Sequence 1, Appli
23	725.5	19.4	1868	1	US-08-658-883B-1	Sequence 1, Appli
24	725.5	19.4	1868	4	US-09-676-610B-26	Sequence 26, Appli
25	724.5	19.4	1593	4	US-09-676-610B-25	Sequence 25, Appli
26	724.5	19.4	5532	2	US-08-475-035-3	Sequence 3, Appli
27	724.5	19.4	5532	4	US-09-676-610B-17	Sequence 17, Appli
28	722.5	19.3	896	6	5200327-2	Patent No. 5200327
29	722.5	19.3	900	1	US-08-318-193-7	Sequence 7, Appli
30	721	19.3	5484	3	US-09-632-580A-3	Sequence 3, Appli
31	721	19.3	5501	1	US-08-484-438-1	Sequence 1, Appli
32	721	19.3	5555	1	US-08-484-438-3	Sequence 3, Appli
33	719	19.2	905	6	5200327-3	Patent No. 5200327
34	719	19.2	909	1	US-08-318-193-9	Sequence 9, Appli
35	714.5	19.1	4905	1	US-07-978-895-3	Sequence 3, Appli
36	714.5	19.1	4905	1	US-08-473-119-3	Sequence 3, Appli
37	714.5	19.1	4905	2	US-08-475-352-3	Sequence 3, Appli
38	714.5	19.1	4905	4	US-09-170-699-3	Sequence 3, Appli
39	714.5	19.1	4975	3	US-09-630-706-3	Sequence 3, Appli
40	710	19.0	1958	4	US-09-570-454-1	Sequence 1, Appli
41	710	19.0	1958	4	US-09-867-521-1	Sequence 1, Appli
42	693	18.5	822	3	US-08-469-318-69	Sequence 69, Appli
43	693	18.5	822	3	US-08-468-609A-69	Sequence 69, Appli
44	693	18.5	822	4	US-08-446-872A-69	Sequence 69, Appli
45	693	18.5	822	4	US-08-762-227A-69	Sequence 69, Appli

ALIGNMENTS

RESULT 1

US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 5518885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: usa
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/229,515A
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 679.537 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLF.....EPVOEGAPPPPAHHHHH 690

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/US09821883/runat_09092004_105126_6623/app_query.fasta_1.3100
-DB=US_Geneseq_29Jan04 -QFWT=fastacp -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN_1_1_1646 @runat_09092004_105126_6623 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:*

1: Geneseq1980as:*
2: Geneseq1990as:*
3: Geneseq2000as:*
4: Geneseq2001as:*
5: Geneseq2002as:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2003cs:*
9: Geneseq2003cs:*
10: Geneseq2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3739	100.0	2070	4 AAD21565	Aad21565 Human HER
2	3473.5	92.9	2091	4 AAD21567	Aad21567 Human HER
3	2974	79.5	1665	4 AAD21564	Aad21564 Human HER
4	2959.5	79.2	1692	4 AAD21566	Aad21566 Human HER
5	2423	64.8	4530	2 AAT01585	Aat01585 Her-2/neu
6	2423	64.8	4530	2 AAT71253	Aat71253 Human HER
7	2423	64.8	4530	3 AA260815	AA260815 Nucleotid
8	2423	64.8	4530	4 AAD19731	Aad19731 Human tyr

9	2423	64.8	4530	6 ABN85585	Abn85585 Human HER
10	2423	64.8	4530	6 ABZ35012	Abz35012 Human gen
11	2423	64.8	4530	6 ABV94128	Abv94128 Breast ca
12	2423	64.8	4530	6 ABK83918	Abk83918 Human cDN
13	2423	64.8	4530	7 ACC50139	Acc50139 Breast ca
14	2423	64.8	4530	7 ABQ83856	Abq83856 Human Her
15	2423	64.8	4530	8 AAD58073	Aad58073 Human C-e
16	2423	64.8	4530	9 ADC09594	Adc09594 Her2/Neu
17	2422	64.8	9274	6 AAD43934	Aad43934 HER-2 tra
18	2422	64.8	9274	6 ABK14057	Abk14057 Human HER
19	2418	64.7	4472	3 AA14812	Aa14812 CDNA enco
20	2418	64.7	4473	2 ABQ76220	Abq76220 Human tum
21	2418	64.7	4473	2 AA231071	Aa231071 HER-2 nuc
22	2418	64.7	4473	6 ABZ34969	Abz34969 Human gen
23	2418	64.7	4473	6 AAD38904	Aad38904 Human Her
24	2418	64.7	4473	7 ACC69999	Acc69999 Human C-e
25	2418	64.7	4473	9 ADC35148	Adc35148 Human bre
26	2418	64.7	4473	9 ADD25483	Add25483 Binding d
27	2415	64.6	9274	4 AAF24297	Aaf24297 HER2 tran
28	2410.5	64.5	3765	9 ADB67620	Adb67620 Human epi
29	2410.5	64.5	3768	3 AAA09455	Aaa09455 Human her
30	2410.5	64.5	3768	6 ABV78168	Abv78168 Human ERB
31	2410.5	64.5	3768	6 ABZ35744	Abz35744 Human ERB
32	2410.5	64.5	3768	6 AAD43935	Aad43935 Human HER
33	2410.5	64.5	3768	6 ABX09987	Abx09987 Human ERB
34	2410.5	64.5	3768	6 AAD43986	Aad43986 Human Her
35	2410.5	64.5	3768	6 ABK14058	Abk14058 Human HER
36	2410.5	64.5	3768	6 ABL91709	Ab191709 Human pol
37	2410.5	64.5	3768	7 ACC57649	Acc57649 Human pro
38	2405.5	64.3	3768	2 AAT40739	Aat40739 HER-2/neu
39	2405.5	64.3	3768	2 AA201912	Aa201912 Human HER
40	2405.5	64.3	3768	5 AA233392	Aa233392 Human HER
41	2405.5	64.3	3768	6 AAD32743	Aad32743 Human Her
42	2405.5	64.3	3768	6 ABK10730	Abk10730 Human Her
43	2405.5	64.3	3768	6 ABQ92250	Abq92250 Human Her
44	2396.5	64.1	4299	2 AAQ46083	Aaq46083 Sequence
45	2393	64.0	4606	9 ADB47370	Adb47370 Human cDN

ALIGNMENTS

RESULT 1

AAD21565

ID AAD21565 standard; DNA; 2070 BP.

XX AC AAD21565;

XX AC AAD21565;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500-hgm-CSF fusion DNA construct.

XX KW Immunostimulatory fusion protein; IPP; antigen component; therapy;

XX KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;

XX KW PAP protein; Ala Arg linker; membrane distal extracellular domain;

XX KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;

XX KW HER-2 protein; granulocyte-macrophage colony stimulating factor;

XX KW HER500-hgm-CSF fusion DNA; ds.

XX OS Homo sapiens.

OS Synthetic.

OS Chimeric.

XX FH Key

XX CDS

XX Location/Qualifiers

1..2070

/*tag= a

/product= "Human HER500-hgm-CSF fusion protein construct

comprising human PAP signal sequence, mature PAP protein,

an Ala Arg linker, human HER-2 signal sequence, mature

HER-2 membrane distal extracellular and intracellular

domains, an Ala Ala linker, a mature human GM-CSF

sequence and a C-terminal tag"

/note= "CDS does not include stop codon"

```

FT XX /partial
FN XX WO200174855-A2.
XX XX 11-OCT-2001.
XX XX 30-MAR-2001; 2001WO-US010515.
XX XX 30-MAR-2000; 2000US-0193504P.
XX XX (DEND-) DENDREON CORP.
XX XX Laus R, Vidovic D, Graddis T;
XX XX WPI; 2001-662965/76.
DR XX P-PSDB; AAE13109.
XX XX
XX XX An immunostimulatory fusion protein comprising the intracellular domain
PT XX of HER-2 and an antigen elicits an immune response to the antigen and is
PT XX useful for the treatment of associated cancer associated.
XX XX
XX XX Example 3; Page 28; 59pp; English.
XX XX
XX XX The invention relates to immunostimulatory fusion proteins (IPF) and
CC XX nucleic acid molecules encoding such proteins. The IPFs comprise a
CC XX polypeptide antigen component and an immunostimulatory component derived
CC XX from the intracellular domain of HER-2 protein which is effective to
CC XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
CC XX immune response to the antigen. IPF or superactivated dendritic cells are
CC XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC XX associated with a particularly antigen. The present sequence is HER500
CC XX hGM-CSF fusion DNA construct which comprises DNA molecules encoding human
CC XX PAP signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
CC XX signal sequence, mature HER-2 membrane distal extracellular and
CC XX intracellular domains, an Ala Ala linker, a mature human granulocyte-
CC XX macrophage colony stimulating factor (GM-CSF) sequence and a C-terminal
CC XX tag
XX XX
XX XX Sequence 2070 BP; 425 A; 707 C; 568 G; 370 T; 0 U; 0 Other;
SQ

```

```

Alignment Scores:
Pred. No.: 1.89e-154 Length: 2070
Score: 3739.00 Matches: 690
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-821-883-2 (1-690) x AAD21565 (1-2070)

Qy 1 MetArgAlaIaProLeuLeuAlaAlaArgAlaAlaSerLeuSerLeuGlyPheIeuPhe 20
Db 1 ATGAGAGTGCACCCCTCTCTGCGCCAGGCGACCAAGCCTTAGCCCTTGGCTTCTGTTT 60
Qy 21 LeuLeuPheThrLeuAspArgSerValLeuAlaIuysGluLeuAlaAtrGlyAlaAla 40
Db 61 CTGCTTTTTTCTGGCTAGACCGAAGTGTACTAGCCAAAGGAGTTGGCGCGCGGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGCACCGGCACACACATGAGCTGGCGCTCCCTGGCCAGTCCCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuThrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGGACATGCTCCCGCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGGAACCTGGAA 240
Qy 81 LeuThrThrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTACCTACCTACCTGCCACCAATGCCAGCTGCTCTCTGCGAGATATCCAGAGAGGTGCAG 300
Qy 101 GlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGCTCATCGCTCACAACCAAGTGAGGAGGTGCCACTGCCACTGCGAGGCTGCGGATT 360

```

```

121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
361 GTGCGAGGCACCCAGCTCTTTTGGAGCACTATGCCCCCTGGCCGTCTAGACAAATGGAGAC 420
141 ProLeuAsnAsnThrThrProValThrGlyValaserProGlyGlyLeuArgGluLeuGln 160
421 CCGCTGAACAATACCAACCCCTGTGCACAGGGGCTCCCCAGAGGCTCGCGGAGCTGCAG 480
161 LeuArgSerLeuThrGluIleLeuLysGlyGlyValLeuLeuGlnArgAsnProGlnLeu 180
481 CTTGGAAGCCTCAGAGATCTTTGAAAGAGGGGTCTTGATCCAGCGGAAACCCCGAGCTC 540
181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu 200
541 TGTACCAGGACACGATTTTGTGAAGAGACATCTTCCACAGAACAAACACAGCTGGCTCTC 600
201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
601 ACACGTAGACACCAACCGCTCTCGGCGCTGCCACCCCTGTTCTCCGATGTGTAGGGC 660
221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
661 TCCGCTGCTGCGGAGAGATTTCTGAGGATTTGTGAGAGCTGACGCGCTGCTGTGTC 720
241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
721 GGTGGCTGTGCGGCTGCAAGGGCCACTGCGCCACTGCTGCTCCATGATGACAGTGTGCT 780
261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
781 GCCGCTGCACGGGCCCAAGCACTCTGACTGCTGCGCTGCGCTCCACTTCAACACAGT 840
281 GlyTleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
841 GGCATCTGTGAGCTGCACCTGCCAGCCCTGCTCACCTACCAACACAGACACGTTTGAAGTCC 900
301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
901 ATGCCCAATCCGAGGCGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGCTCCCTAC 960
321 AsnTyrLeuSerThrAspValGlySerGlyAlaGlyMetValHisHisArgHisArg 340
961 AACTACCTTTTCTACGAGCTGGGATCGGGCGCTGGGGGCATGGTCCACACAGCACCGC 1020
341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
1021 AGCTCATCTACAGGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAG 1080
361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
1081 GAGGCCCCCAGGTCTCCACTGGCCACCTCCGAGGGGCTGGCTCCGATGTATTGTATGGT 1140
381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
1141 GACCTGGGAATGGGGCAGCAAGGGCTGCAAGAGCTCCCAACACACATGACCCCGAGCCT 1200
401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
1201 CTACAGCGGTACATGATGAGGACCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
1261 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
1321 CCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
1381 AGGCGCAAGACTCTCTCCCGAGGAGAAATGGGGTGTCTCAAGAGACGTTTTTGGCTTTGGG 1440

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein.- nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 7392.31 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09821883/runat.09092004.105127.6631/app.query.fasta.1.3100
-DB=genEmbl -QWTF=frascap -SUFFIX=rge -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883 @CGN 1.1 16795 @runat.09092004.105127.6631 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.ats.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.ats.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3739	100.0	2070	6	AX268286 Sequence
2	3473.5	92.9	2091	6	AX268288 Sequence
3	2974	79.5	1665	6	AX268285 Sequence
4	2959.5	79.2	1692	6	AX268287 Sequence
5	2423	64.8	4530	6	I21124 Sequence 9
6	2423	64.8	4530	6	I59745 Sequence 9
7	2423	64.8	4530	6	AR202597 Sequence
8	2423	64.8	4530	6	AR283481 Sequence
9	2423	64.8	4530	6	AR344811 Sequence
10	2423	64.8	4530	6	AX282577 Sequence
11	2423	64.8	4530	6	AX587649 Sequence
12	2423	64.8	4530	6	AX644071 Sequence
13	2423	64.8	4530	6	AX771418 Sequence
14	2423	64.8	4530	6	BD005474 Cellular
15	2423	64.8	4530	9	HUMHER2A
16	2422	64.8	9274	6	AR409602 Sequence
17	2422	64.8	9274	6	AX060703 Sequence
18	2418	64.7	4473	6	AR080259 Sequence
19	2418	64.7	4473	6	AR167390 Sequence
20	2418	64.7	4473	6	AR392088 Sequence
21	2418	64.7	4473	9	HSERB2R
22	2410.5	64.5	3768	6	AR409603 Sequence
23	2410.5	64.5	3768	6	AX060704 Sequence
24	2410.5	64.5	3768	6	AX467229 Sequence
25	2410.5	64.5	3768	6	AX481438 Sequence
26	2410.5	64.5	3768	6	BD224136 Novel met
27	2405.5	64.3	3768	6	AR034479 Sequence
28	2405.5	64.3	3768	6	BD267514 HER-2/neu
29	2405.5	64.3	3768	6	AX201817 Sequence
30	2405.5	64.3	3768	6	AX380923 Sequence
31	2405.5	64.3	3768	6	AX384604 Sequence
32	2405.5	64.3	3768	6	AX465456 Sequence
33	2334.5	62.4	3678	6	AX505114 Sequence
34	2258.5	60.4	2385	6	AR082744 Sequence
35	2258.5	60.4	2385	6	AR099963 Sequence
36	2258.5	60.4	2385	6	AR143949 Sequence
37	2188.5	58.5	1437	6	AX268289 Sequence
38	2130	57.0	2781	6	AX380944 Sequence
39	2129	56.9	2763	6	AX380942 Sequence
40	2124.5	56.8	3780	4	AB008451 Canis fam
41	1974	52.8	4506	10	AK129487 Mus muscu
42	1972	52.7	4727	10	AY116182 Rattus no
43	1968	52.6	3955	6	BD267515 HER-2/neu
44	1968	52.6	3955	6	I21129 Sequence 14
45	1968	52.6	3955	6	I59750 Sequence 14

ALIGNMENTS

AX268286
LOCUS AX268286 2070 bp DNA linear PAT 29-OCT-2001
DEFINITION Sequence 7 from Patent WO0174855.
ACCESSION AX268286
VERSION AX268286.1 GI:16541538
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Laus, R., Vidovic, D. and Graddis, T.
TITLE Compositions and methods for dendritic cell-based immunotherapy
JOURNAL Patent: WO 0174855-A 7 11-OCT-2001;
DENDREON CORPORATION (US)
FEATURES
source Location/Qualifiers
1..2070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="HER500-hgm-CSF construct"

ORIGIN
Alignment Scores:
Pred. No.: 7,45e-129 Length: 2070
Score: 3739,00 Matches: 690
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-821-883-2 (1-690) x AX268286 (1-2070)

Qy 1 MethArgAlaProLeuLeuLeuAlaArgAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGCACCCCTCTCTGGCCAGGACAGCAAGCCCTTAGCCCTTGGCTTCTGTTT 60
Qy 21 LeuLeuPhePheTTPLeuAspArgSerValLeuAlaLysGlyLeuLeuAlaArgGlyAlaAla 40
Db 61 CTGCTTTTTTCTGGCTAGACCGAAGTGTACTACCAAGAGGTTGGCGCGGCGGCGCG 120
Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCGAGACC 180
Qy 61 HisLeuAspMetLeuArgHisLeuTyrglnClyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGGACATGTCTCCGCCACCTCTACACGGGCTGCCAGGTGGTGCAGGGAAACCTGGAA 240
Qy 81 LeuThrTyrrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGln 100
Db 241 CTACACTACCTGCACCAATGCCAGCTGTCTCTCTCCGAGGATATCCAGGAGTGCAG 300
Qy 101 GlyTyrrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgIle 120
Db 301 GGCTACGTGTCTATCGCTCAACCAAGTGAAGGAGGTCCCACTGCAGAGGCTCGGATT 360
Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCCACCCAGCTCTTTGAGGACAACTATCTCCCTGGCGGTGTAGCAATGGAGAC 420
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCGCTGAACATACACCCCTGTCTACAGGGGCTCCCAAGAGGAGGCTTGGGGAGTGCAG 480
Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyGlyValLeuIleGlnArgAsnProGlnLeu 180
Db 481 CTTGGAAGCTTCACAGAGATCTTGAAGAGGGGTCTTGATCCAGCGCAACCCCAAGCTC 540
Qy 181 CysTyrrGlnAspThrIleLeuTTPLeuAspIlePheHisLysAsnAsnGlnLeuLeuLeu 200
Db 541 TGCTACAGGACACGATTTTGTGGAAGGACATCTCCACAGAAACACCAAGCTGGCTCTC 600
Qy 201 ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220

Db 601 ACATGATAGACACCAACCGCTCTCGGGCTCGCCACCCCTGTTCTCCGATGTGTAAGGCG 660
Qy 221 SerArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGCTGGGAGAGAGTTCTGAGGATTTGTGAGAGCCTGAGCGCCTGCTGTGCTG 720
Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGCCCTGCAAGGGGCCACTGCGCCACTGCTGCTGCTCATGAGCATGTGCT 780
Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTGCACGGGCCCAAGCACTCTGACTGCTGCTGCTGCTGCCCTCAACACAGT 840
Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGTGTGACTGCCAGCCCTGGTACCTACCAACACAGACACAGATTTGAGTCC 900
Qy 301 MetProAsnProGluGlyArgTyrrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGGGCCGGTATACATTCGGGCCACAGCTGTGTGACTGCTGCTGCTTAC 960
Qy 321 AsnTyrrLeuSerThrAspValGlySerGlyValaGlyGlyMetValHisHisArgHisArg 340
Db 961 AACTACCTTTCTACGGAGCTGGGATCGGCGCTGGGGGCATGGTCCACACAGCACCGC 1020
Qy 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
Db 1021 AGCTCATCTACAGAGTGGCGGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAAGAG 1080
Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCCCCAGAGTCTCCACTGGCACCCCTCCGAGGGGCTGGCTCCGATGATTTGATGGT 1140
Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACCTGGGAATGGGGCAGCAAGGGCTGCAAAAGCTCCCAACACATGATGATGCTGCT 1200
Qy 401 LeuGlnArgTyrrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrrVal 420
Db 1201 CTACAGCGGTACAGTGAAGACCCACAGTACCCCTGCCCCCTCTGAGATGATGCTGCT 1260
Qy 421 AlaProLeuThrCysSerProGlnProGluTyrrValAsnGlnProAspValArgProGln 440
Db 1261 GCCCCCTGACTGACGCCCCAGCTGATATGTGAACACAGCAGATGTTTCGGCCCCAG 1320
Qy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
Db 1321 CCCCCTTCGCCCGAGAGGGCCCTCTGCTGCTGCCGACCTGCTGGTGGCACTCTCGAA 1380
Qy 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
Db 1381 AGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTCTGCAAGAGCTTTTTCCTTTGGG 1440
Qy 481 GlyAlaValGluAsnProGluTyrrLeuThrProGlnGlyGlyAlaAlaProGlnProHis 500
Db 1441 GGTGCGGTGGAGAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCCCTCAGCCCCAC 1500
Qy 501 ProProProAlaPheSerProAlaPheAspAsnLeuTyrrTyrrAspGlnAspProPro 520
Db 1501 CCTCCTCTCCTTCAGCCCCAGCCCTTCGACAAACCTCTATTACTGGGACAGGACCCACCA 1560
Qy 521 GluArgGlyAlaProProSerThrPheLysGlyThrProThrAlaGluAsnProGluTyr 540
Db 1561 GAGCGGGGGCTCCACCCAGCAGCTTCAAGAGGACACTACGGCAGAGAACCCACAGATGAC 1620
Qy 541 LeuGlyLeuAspValProAlaAlaProAlaArgSerProSerProSerThrGlnPro 560
Db 1621 CTGGGTCTGAGCTGTCAGCGGGCGCACCCCGCGCTCGCCCGAGCCCCCAGCACACAGCCC 1680
Qy 561 TrpGluHisValAsnAlaIleGlnGluAlaArgArgLeuLeuAsnLeuSerArgAspThr 580

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 68.5595 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLAARASLSGLF.....EPVOEGAPPPAAAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2124.5	56.8	1259	6	O18735 canis famil
2	1951	52.2	1259	11	Q8K3F9
3	1609.5	43.0	419	4	Q9UKY9
4	1376.5	36.8	711	11	Q8OV89
5	1005.5	26.9	881	11	Q8C0E7
6	969	25.9	367	11	Q8R2X1
7	906	24.2	165	4	Q14256
8	900.5	24.1	412	4	Q8WYV0
9	861.5	23.0	431	13	Q7SYI9
10	754.5	20.2	527	13	Q90836
11	723	19.3	149	6	Q8BG66
12	717.5	19.2	643	11	Q9ERV6
13	717.5	19.2	655	11	Q9WVP5
14	717.5	19.2	1210	11	Q9EP98
15	714.5	19.1	1209	11	Q9QX70
16	708.5	18.9	1209	6	Q8MIL8

17	704.5	18.8	478	11	Q9ESE0
18	703	18.8	1191	13	Q7SZF7
19	697	18.6	331	4	Q9BUD7
20	664.5	17.8	1165	13	Q9YH40
21	657.5	17.6	1305	13	Q8AW81
22	654	17.5	144	6	Q9GL44
23	641.5	17.2	599	13	Q9PSH2
24	634	17.0	144	6	Q86SY5
25	608	16.3	1328	13	P79754
26	556	14.9	1377	5	Q8MLW0
27	545	14.6	144	6	Q9MYK4
28	543	14.5	1433	5	Q9BIH9
29	537	14.4	1322	5	Q86NZ2
30	527	14.1	146	6	Q8WNL7
31	515	13.8	152	6	Q9SL10
32	473	12.7	141	6	Q7YRF7
33	427	11.4	141	11	Q8VH40
34	407	10.9	1137	13	Q9WGF6
35	398	10.6	150	6	Q9BG64
36	395	10.6	151	6	Q9BG65
37	382.5	10.2	145	11	Q8CFB5
38	377	10.1	1368	5	Q23821
39	362	9.7	138	11	Q9UJ91
40	341	9.1	365	5	Q26589
41	341	9.1	1717	5	Q26586
42	331	8.9	334	5	Q26567
43	331	8.9	342	5	Q26568
44	323.5	8.7	1193	5	Q9Y1X8
45	305	8.2	1564	5	Q86MD7

ALIGNMENTS

RESULT 1

O18735 PRELIMINARY; PRT; 1259 AA.
AC O18735;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Erbb-2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_taxid=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Yokota H.;
RT "CDNA cloning of erbb-2 from canine mammary gland.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008451; BAA23127.1; -
DR HSSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . .; IEA.
DR InterPro; IPR000494; EGFR_L_domain.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR InterPro; IPR004019; YLP motif.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF01030; Recep_L_domain; 2.
DR Pfam; PF02757; YLP; 2.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 12.6656 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLF.....EPVQEGAPPPPAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2405.5	64.3	1255	1 ERB2_HUMAN	P04626 homo sapien
2	1947.5	52.1	1257	1 ERB2_RAT	P06494 rattus norv
3	1935	51.8	1254	1 ERB2_MESAU	Q60553 mesocricetu
4	755.5	20.2	703	1 EGFR_CHICK	P13387 gallus gall
5	721	19.3	1210	1 EGFR_HUMAN	P00533 homo sapien
6	721	19.3	1308	1 ERB4_HUMAN	P01303 homo sapien
7	717.5	19.2	1210	1 EGFR_MOUSE	Q01279 mus musculu
8	717	19.2	1308	1 ERB4_RAT	Q62956 rattus norv
9	697	18.6	1342	1 ERB3_HUMAN	P21860 homo sapien
10	674	18.0	144	1 CSF2_HUMAN	P04141 homo sapien
11	668	17.9	1339	1 ERB3_RAT	Q62799 rattus norv
12	655.5	17.5	1167	1 XMRK_XIPMA	P13388 xiphophorus
13	549	14.7	144	1 CSF2_SHEEP	P28773 ovis aries
14	536	14.3	1426	1 EGFR_DROME	P04412 drosophila
15	529	14.1	144	1 CSF2_CEREL	P51748 cervus elap
16	488	13.1	144	1 CSF2_PIG	Q29118 sus scrofa
17	481.5	12.9	143	1 CSF2_BOVIN	P11052 bos taurus
18	476.5	12.7	144	1 CSF2_CANFA	P48749 canis famil
19	452	12.1	144	1 CSF2_FELCA	Q62757 felis silve
20	441	11.8	127	1 CSF2_RAT	P48750 rattus norv
21	435.5	11.6	140	1 CSF2_CAVPO	Q60481 cavia porce
22	404.5	10.8	1367	1 LTJ3_CAEEL	P24348 caenorhabdi
23	367	9.8	141	1 CSF2_MOUSE	P01587 mus musculu
24	317.5	8.5	1363	1 ILPR_BRALA	Q02466 branchiosteo
25	293	7.8	147	1 ERB2_MOUSE	P70424 mus musculu
26	273.5	7.3	245	1 HTK7_HYDAT	Q25197 hydra atten
27	270.5	7.2	2146	1 INSR_DROME	P06208 drosophila
28	256.5	6.9	1382	1 INSR_HUMAN	P06213 homo sapien
29	252.5	6.8	1372	1 INSR_MOUSE	P15208 mus musculu
30	252	6.7	1607	1 MIPR_LYMST	Q25410 lymphaea sca
31	251.5	6.7	1300	1 IRR_MOUSE	Q9wt14 mus musculu
32	249	6.7	1383	1 INSR_RAT	P15127 rattus norv
33	248.5	6.6	1297	1 IRR_HUMAN	P14616 homo sapien

34	246.5	6.6	1300	1 IRR_CAVPO	P14617 cavia porce
35	246	6.6	1367	1 IGIR_HUMAN	P08069 homo sapien
36	235.5	6.3	1370	1 IGIR_RAT	P24062 rattus norv
37	233.5	6.2	1373	1 IGIR_MOUSE	Q60751 mus musculu
38	228.5	6.1	581	1 IRR_RAT	Q64716 rattus norv
39	207	5.5	1390	1 INSR_AEDAE	Q93105 aedes aegypt
40	178	4.8	2715	1 MLL4_HUMAN	Q9um06 homo sapien
41	168	4.5	386	1 PPAP_HUMAN	P15309 homo sapien
42	166	4.4	707	1 SPFQ_HUMAN	P23246 homo sapien
43	166	4.4	5262	1 MLL2_HUMAN	O14686 homo sapien
44	162	4.3	1321	1 IRS2_MOUSE	P81122 mus musculu
45	160	4.3	830	1 SREC_HUMAN	Q14162 homo sapien

ALIGNMENTS

RESULT 1
ERB2_HUMAN
ID ERB2_HUMAN STANDARD; PRT; 1255 AA.
AC P04626;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)
DE (p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell surface receptor HER2) (MLN 19).
DE ERBB2 OR HER2 OR NGL OR NEU.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86119663; PubMed=3003577;
RA Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N., Saito T., Toyoshima K.;
RT "Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth factor receptor.";
RL Nature 319:230-234(1986).
[2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-1170.
RX MEDLINE=86070181; PubMed=2999974;
RA Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A., McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J., Francke U., Levinson A., Ullrich A.;
RT "Tyrosine kinase receptor with extensive homology to EGF receptor shares chromosomal location with neu oncogene.";
RL Science 230:1132-1139(1985).
[3]
RP SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.
RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W., Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D., Schackwitz W.S., Sharwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 737-1031 FROM N.A.
RX MEDLINE=86016729; PubMed=2955967;
RA Semba K., Kamata N., Toyoshima K., Yamamoto T.;
RT "A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epidermal growth factor-receptor gene and is amplified in a human salivary gland adenocarcinoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).
[5]
RP VARIANTS VAL-654 AND VAL-655.
RX MEDLINE=93194196; PubMed=8095488;
RA Ehsani A., Low J., Wallace R.B., Wu A.M.;
RT "Characterization of a new allele of the human ERBB2 gene by allele-specific competition hybridization.";
RL Genomics 15:426-429(1993).
CC -!- FUNCTION: Essential component of a neuroregulin-receptor complex, although neuroregulins do not interact with it alone. GP30 is a potential ligand for this receptor. Not activated by EGF, TGF-

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 22.0271 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-2
Perfect score: 3739
Sequence: 1 MRAAPLLARAASLSGLFLF.....BPVOEGAPPPAAAHHHHHH 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405.5	64.3	1255	1 A24571	protein-tyrosine k
2	1950.5	52.2	1260	1 TVRTNU	protein-tyrosine k
3	1935	51.8	1254	2 I48161	p-185 precursor -
4	755.5	20.2	1223	1 TVCHLV	epidermal growth f
5	754.5	20.2	527	2 A42032	epidermal growth f
6	721	19.3	1210	1 GQHUE	epidermal growth f
7	721	19.3	1308	2 A47253	epidermal growth f
8	717.5	19.2	1210	2 A53183	epidermal growth f
9	714.5	19.1	644	2 A36325	epidermal growth f
10	697	18.6	1342	2 A36223	kinase-related tra
11	674	18.0	144	1 FQHUGM	granulocyte-macrop
12	662	17.7	1339	2 JCA387	epidermal growth f
13	655.5	17.5	1166	1 S06142	protein-tyrosine k
14	549	14.7	144	2 JH0469	granulocyte-macrop
15	545	14.6	144	1 A61632	granulocyte-macrop
16	536	14.3	843	2 A27131	epidermal growth f
17	481.5	12.9	143	1 FQBOGM	granulocyte-macrop
18	476.5	12.7	144	2 A44936	granulocyte-macrop
19	441	11.8	127	2 I46269	granulocyte-macrop
20	404.5	10.8	1323	2 E88257	protein let-23 (im
21	404.5	10.8	1374	2 S70712	protein-tyrosine k
22	390.5	10.4	1369	2 S70713	protein-tyrosine k
23	375.5	10.0	1330	1 GQFPE	epidermal growth f
24	374	10.0	153	1 FQMSGM	granulocyte-macrop
25	341	9.1	1767	2 D45558	epidermal growth f
26	341	9.1	361	1 A45558	epidermal growth f
27	331	8.9	333	2 B45558	epidermal growth f
28	331	8.9	342	2 C45558	epidermal growth f
29	317.5	8.5	1363	2 T43220	insulin-like growth

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N:Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB2
C:Species: Homo sapiens (man)
C:Date: 25-Oct-1987 #sequence_revision 06-Dec-1996 #text_change 11-Jun-1999
C:Accession: A24571; A25491; A44188; B44188; I59509; I57622
R:Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A:Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f
A:Reference number: A24571; MUID:86118663; PMID:3003577
A:Accession: A24571
A:Molecule type: mRNA
A:Residues: 1-1255 <YAM>
A:Cross-references: GB:X03363; NID:G31197; PIDN:CAA27060.1; PID:G31198
R:Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A:Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm
A:Reference number: A25491; MUID:86016729; PMID:2995967
A:Accession: A25491
A:Molecule type: DNA
A:Residues: 737-1031 <SEM>
A:Cross-references: GB:M11767; NID:G182163; PIDN:AAA35808.1; PID:G553282
R:Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, E.
Science 230, 1132-1139, 1985
A:Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos
A:Reference number: A44188; MUID:86070181; PMID:2999974
A:Accession: A44188
A:Molecule type: DNA
A:Residues: 740-910 <COUL>
A:Cross-references: GB:M12036; NID:G183988; PIDN:AAA35978.1; PID:G183989
A:Accession: B44188
A:Molecule type: mRNA
A:Residues: 1-517, 'RALL', 522, 'S', 524-554, 'V', 656-1169, 'A', 1171-1255 <COU2>
A:Cross-references: GB:M11730; NID:G183986
R:King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A:Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A:Reference number: I59509; MUID:85272597; PMID:2992089
A:Accession: I59509
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 832-909 <REX>
A:Cross-references: GB:L29395; NID:G459807; PIDN:AAA35809.1; PID:G459808
R:Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A:Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A:Reference number: I57622; MUID:87286898; PMID:3039351
A:Accession: I57622
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-191 <TAL>

A;Cross-references: GB:M16792; NID:g183983; PIDN:AAA58637.1; PID:g553332
C;Comment: Amplification and overexpression of this erbB-related gene occurs in about 30
C;Genetics:
A;Gene: GDB:ERBB2; NGL; NEU; HER-2
A;Cross-references: GDB:120613; OMIM:164870
A;Map position: 17q21.1-17q21.1
A;Introns: 25/1; 75/3; 147/1; 883/3
A;Note: the list of introns is incomplete
C;Function:
A;Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
inase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;12-1255/Product: protein-tyrosine kinase erbB2 #status predicted <MAT>
F;122-653/Domain: extracellular #status predicted <EXT>
F;70-304/Domain: EGF receptor extracellular domain repeat <EE1>
F;395-605/Domain: EGF receptor extracellular domain repeat <EE2>
F;654-675/Domain: transmembrane #status predicted <TM>
F;676-1255/Domain: intracellular #status predicted <INT>
F;718-983/Domain: intracellular #status predicted <KIN>
F;726-734/Region: protein kinase ATP-binding motif
F;68,124,187,259,530,571,629/Binding site: carboxydrate (Asn) (covalent) #status predicted
F;686/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;753/Active site: Lys #status predicted
F;1139,1221,1222,1248/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation)

Query Match 64.3%; Score 2405.5; DB 1; Length 1255;
Best Local Similarity 41.1%; Pred. No. 1.4e-125;
Matches 509; Conservative 0; Mismatches 3; Indels 727; Gaps 1;

Qy 35 LARGASTOVCTGDMKLRLPASPTHLDMLRHLYGCGVQGNLETLVLPNASLSPLQ 94
Db 16 LPPGAASQVCTGDMKLRLPASPTHLDMLRHLYGCGVQGNLETLVLPNASLSPLQ 75
Qy 95 DIQEVQGVLIHQNQVQVPLQRLIRVGTQQLFEDNYALALDNGDPLNNTPVTGASPG 154
Db 76 DIQEVQGVLIHQNQVQVPLQRLIRVGTQQLFEDNYALALDNGDPLNNTPVTGASPG 135
Qy 155 GLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALTLIDNRSRACHPC 214
Db 136 GLRELQRLSLTEILKGGVLIQRNPOLCYQDTILWKDIFHKNNQALALTLIDNRSRACHPC 195
Qy 215 SPMCKGSRGWESSEDCOSLRTVCAGGACRCKGPLEPTDCHEOCACAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRGWESSEDCOSLRTVCAGGACRCKGPLEPTDCHEOCACAGCTGPKHSDCLAC 255
Qy 275 LHFHNSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGS----- 329
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNPGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
Qy 330 ----- 329
Db 316 PLHNQEVTAEDGTORCEKSKPCARVCYGLGWEHLREVRVTSANIQFAGCKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPSFDGDPASNTAPLOEQVLFETLBEITGYLYISAWPDSLPDLVSFQNLQVIRGRI 435
Qy 330 ----- 329
Db 436 LHNGAYSILTQLOGISLWGLRSLRSGSLALIHNTLHCFVHTVPWDQLFRNPQHALLH 495
Qy 330 ----- 329
Db 496 TANRPEDECVGEGLAGHOLCARGHCWGPPTQCVNCSQFLRQGCVECRVLQGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADOCVCAHYKDPFVCARCPGSKVKKPDLSPYIWK 615
Qy 330 ----- 329

Db 616 FPDEGACQPCPINCTHSCVDLDDKGCPAEQRASPLTIIISAVVGILLVVLGVVFGILI 675
Qy 330 ----- 329
Db 676 KRRQOKIRKYMRRELOETELVEPLTPSGAMPNOAQMRILKETELRKVKVLGSGAFGVY 735
Qy 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTSPKANKEILDEAYVMAGVGSPPYVSRLLGICLTSTVQ 795
Qy 330 ----- 329
Db 796 LVTQLMPVGCILLDHVRENRRGLSGQDLLNMCWQIAKGMSYLEDVRLVHRDLAARNVLKVS 855
Qy 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGKVPKIKWMALESILARRRTHOSDVMSYGVTVWEL 915
Qy 330 ----- 329
Db 916 MTFCAKPYDGIPIAREIPDLLEKGERLPQPPCTTIDVYMVKWIDSECRPRELVS 975
Qy 330 ----- 329
Db 976 FSRWARDPQRFVITQNEEDLGPASPLDSTFYRSLLLEDDMGDLVDAEEYLVPOQGFCDP 1035
Qy 330 --GAGGMVHRHRSSTSSRSGGDLTLGLEPSEEPAPSPAPSGAGSDVFDGLMGAA 387
Db 1036 APGAGGMVHRHRSSTSSRSGGDLTLGLEPSEEPAPSPAPSGAGSDVFDGLMGAA 1095
Qy 388 KGLQSLTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREG 447
Db 1096 KGLQSLTHPSPLOQRYSEDPVPLPSETDGYVAPLTCSPQPEYVNPDPVRPQPPSPREG 1155
Qy 448 PLPAARPAAGATLESAKTLSPGKNGVVKDVFAFGAVENPEYLTPOGGGAAPQHPHPPAFSP 507
Db 1156 PLPAARPAAGATLESAKTLSPGKNGVVKDVFAFGAVENPEYLTPOGGGAAPQHPHPPAFSP 1215
Qy 508 AFDNLYWDQDPPRGGAPPSTFKGTPTAENPEYLGLDVP 546
Db 1216 AFDNLYWDQDPPRGGAPPSTFKGTPTAENPEYLGLDVP 1254

RESULT 2
TVRTNU
protein-tyrosine kinase (EC 2.7.1.112) neu precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 11-Jun-1999
C;Accession: A24562; A61204
R;Bargmann, C.I.; Hung, M.C.; Weinberg, R.A.
Nature 319, 226-230, 1986
A;Title: The neu oncogene encodes an epidermal growth factor receptor-related protein.
A;Reference number: A24562; MUID:86118662; PMID:3945311
A;Accession: A24562
A;Molecule type: mRNA
A;Residues: 1-1260 <BAR>
A;Cross-references: EMBL:X03362; NID:g56745; PIDN:CAA27059.1; PID:g56746
R;Masui, T.; Mann, A.M.; Macatee, T.L.; Garland, E.M.; Okamura, T.; Smith, R.A.; Cohen,
Carcinogenesis 12, 1975-1978, 1991
A;Title: Direct DNA sequencing of the rat neu oncogene transmembrane domain reveals no m
2-thiazolyl]formamide or N-methyl-N-nitrosourea.
A;Reference number: A61204; MUID:92035293; PMID:1682063
A;Accession: A61204
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 637-663,'V',665-702 <MAS>
A;Note: authors translated the codon GCA for residue 25 as Val
C;Genetics:
A;Gene: neu
C;Superfamily: epidermal growth factor receptor; protein kinase homology
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; phosphoprotein; phospho
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1260/Product: protein-tyrosine kinase neu #status predicted <MAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 72.9649 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARASLSGLFLF.....EPVQEGAPPPAANHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/FCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3739	100.0	690	US-09-821-883-2	Sequence 2, Appli
2	3473.5	92.9	697	US-09-821-883-4	Sequence 4, Appli
3	2972	79.5	555	US-09-821-883-1	Sequence 1, Appli
4	2957.5	79.1	564	US-09-821-883-3	Sequence 3, Appli
5	2573.5	68.8	919	US-09-854-356-6	Sequence 6, Appli
6	2410.5	64.5	1255	US-09-811-123-9	Sequence 9, Appli
7	2410.5	64.5	1255	US-09-811-115-3	Sequence 3, Appli
8	2410.5	64.5	1255	US-09-769-508-2	Sequence 2, Appli
9	2410.5	64.5	1255	US-09-984-092-4	Sequence 4, Appli
10	2410.5	64.5	1255	US-10-177-293-126	Sequence 126, App
11	2410.5	64.5	1255	US-10-207-498-6	Sequence 6, Appli
12	2410.5	64.5	1255	US-10-338-730-2	Sequence 2, Appli
13	2410.5	64.5	1255	US-10-322-892-4	Sequence 2, Appli
14	2410.5	64.5	1255	US-10-272-437A-28	Sequence 28, Appli
15	2410.5	64.5	1255	US-10-117-937-594	Sequence 594, App

16	2410.5	64.5	1255	15	US-10-435-696-36	Sequence 36, Appli
17	2410.5	64.5	1255	16	US-10-441-779C-4	Sequence 4, Appli
18	2410.5	64.5	1255	16	US-10-734-1564-126	Sequence 126, App
19	2405.5	64.3	1255	9	US-09-854-356-1	Sequence 1, Appli
20	2405.5	64.3	1255	9	US-09-930-125-2	Sequence 2, Appli
21	2405.5	64.3	1255	10	US-09-441-411-6	Sequence 6, Appli
22	2405.5	64.3	1255	12	US-10-469-162-3	Sequence 3, Appli
23	2405.5	64.3	1255	12	US-10-253-286-553	Sequence 553, App
24	2405.5	64.3	1255	12	US-09-765-973-2	Sequence 2, Appli
25	2405.5	64.3	1255	12	US-10-418-027-3	Sequence 3, Appli
26	2405.5	64.3	1255	14	US-10-207-655-45	Sequence 45, Appli
27	2405.5	64.3	1255	14	US-10-313-644-2	Sequence 2, Appli
28	2405.5	64.3	1255	15	US-10-394-322A-17	Sequence 17, Appli
29	2405.5	64.3	1255	15	US-10-245-871-553	Sequence 553, App
30	2405.5	64.3	1255	15	US-10-149-138-4641	Sequence 4641, App
31	2405.5	64.3	1255	16	US-10-647-005-68	Sequence 68, Appli
32	2405.5	64.3	1255	16	US-10-149-138-4641	Sequence 4641, App
33	2398.5	64.1	1253	14	US-10-146-473-72	Sequence 72, Appli
34	2188.5	58.5	479	9	US-09-821-883-5	Sequence 5, Appli
35	1961	52.4	1256	9	US-09-854-356-14	Sequence 14, Appli
36	1950.5	52.2	1260	9	US-09-870-759-118	Sequence 118, App
37	1950.5	52.2	1260	10	US-09-751-708A-118	Sequence 118, App
38	1948	52.1	1256	9	US-09-854-356-2	Sequence 2, Appli
39	1612	43.1	645	9	US-09-921-161-1	Sequence 1, Appli
40	1612	43.1	645	14	US-10-268-501-13	Sequence 13, Appli
41	1612	43.1	645	15	US-10-608-626-13	Sequence 13, Appli
42	1612	43.1	653	9	US-09-854-356-3	Sequence 3, Appli
43	1612	43.1	685	15	US-10-412-804A-4	Sequence 4, Appli
44	1612	43.1	690	15	US-10-412-804A-11	Sequence 11, Appli
45	1612	43.1	712	9	US-09-854-356-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-821-883-2
; Sequence 2, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-hGM-CSF construct
US-09-821-883-2

Query Match	100.0%	Score 3739;	DB 9;	Length 690;
Best Local Similarity	100.0%	Pred. No. 1.8e-244;		
Matches 690;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MRAAPLLARASLSGLFLFLLFWLDRSLAKELARGAASSTQVCTGDMKRLPASPET	60	
Db	1	MRAAPLLARASLSGLFLFLLFWLDRSLAKELARGAASSTQVCTGDMKRLPASPET	60	
Qy	61	HLDMLRHLYQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIANQVQVPLQRLRI	120	
Db	61	HLDMLRHLYQCQVQVQGNLEITYLPTNASLSFLQDIQEVQGVLIANQVQVPLQRLRI	120	
Qy	121	VRGTQLFEDNYALAVLDNGDPPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIORNPOL	180	

```
Db 121 VRGTLFEDNYALAVLNGDPLNNTTPTVGASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYILSTDVSGAGGVMVHRRHRSSTSGGGDLTLGLPSEE 360
Db 301 MPNPEGRTYTFGASCVTACPNYILSTDVSGAGGVMVHRRHRSSTSGGGDLTLGLPSEE 360
Qy 361 EAPRSLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYV 420
Db 361 EAPRSLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPVPLPSETDGYV 420
Qy 421 APLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNGVVKDVFAG 480
Db 421 APLTCSPOPEYVNPQDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNGVVKDVFAG 480
Qy 481 GAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Qy 541 LGLDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVISEMFDLQEP 600
Db 541 LGLDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVISEMFDLQEP 600
Qy 601 TCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITTFESFKENLKD 660
Db 601 TCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITTFESFKENLKD 660
Qy 661 FLLVIPDCWEPVQEGAPPPPPAAAHHHHHH 690
Db 661 FLLVIPDCWEPVQEGAPPPPPAAAHHHHHH 690

RESULT 2
US-09-821-883-4
; Sequence 4, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500*-rGM-CSF construct
US-09-821-883-4

Query Match 92.9%; Score 3473.5; DB 9; Length 697;
Best Local Similarity 91.7%; Pred. No. 1.7e-226;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;

Qy 1 MRAAPLLIARAASISLGLFLLFFWLDRSLVAKELARGAASQTGCTGDMKRLPASPSET 60
Db 1 MRAAPLLIARAASISLGLFLLFFWLDRSLVAKELARGAASQTGCTGDMKRLPASPSET 60
```

```
Qy 61 HLDMLRHLQCVQVQGNLELTYLPTNASLSFLQDIOEVQGYVLIAHNQVRQVPLQRLRI 120
Db 61 HLDMLRHLQCVQVQGNLELTYLPTNASLSFLQDIOEVQGYVLIAHNQVRQVPLQRLRI 120
Qy 121 VRGTLFEDNYALAVLNGDPLNNTTPTVGASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Db 121 VRGTLFEDNYALAVLNGDPLNNTTPTVGASPGGLRELQLRSLEILKGGVLIQRNPQL 180
Qy 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGESSEDCQSILTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALALIDTNRSRACHPCSPMCKGRCWGESSEDCQSILTRTVCA 240
Qy 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
Qy 301 MPNPEGRTYTFGASCVTACPNYILSTDVSGS-----GAGGMVHRRHRSSTSGGGDL 351
Db 301 MPNPEGRTYTFGASCVTACPNYILSTDVSGSIIINFEKLGAGGMVHRRHRSSTSGGGDL 360
Qy 352 TLGLPSEEEAPRSLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPVPL 411
Db 361 TLGLPSEEEAPRSLAPSGAGSDVDFDGLGMAAGLQSLPHTDPSPLQRYSEDPVPL 420
Qy 412 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 471
Db 421 LPSETDGYVAPLTCSPQPEYVNPQDVRPQPPSPREGPLPAARPAAGATLERAKTILSPGKNG 480
Qy 472 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 531
Db 481 VVKDVFAGGAVENPEYLTPOGGAAPQHPHPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 540
Qy 532 TPTAENPEYLGIDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVI 591
Db 541 TPTAENPEYLGIDVPAAPARSPSPSTQPMHEVNAIOEARLLNLSRDTAAEMNETVEVI 600
Qy 592 SEMEDLQEPCLQTRLELYKQGLRGSITKLKGPLTMASHYKQHCPTPTSCATQIITTF 651
Db 601 SNEFSIORPTCVQTRLXLYKQGLRGNLTKLNGALTMTIASHYQTNCPTPETDCBIEVTTF 660
Qy 652 ESPKENLKDIFLLVIPDCWEPVQEGAPPPPPAAAHHHHHH 690
Db 661 EDFIKNLKGLFDIPDFDCWKPQKGA PPPP--AHHHHHH 697

RESULT 3
US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; PRIOR FILING DATE: 2001-03-30
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 79.5%; Score 2972; DB 9; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e-192;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 20.0998 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAASLSGLFLF.....EPVOEGAPPPAAAAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/ECTUS_COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410.5	64.5	1255	4	US-09-527-487-2
2	2410.5	64.5	1255	4	US-09-811-115-3
3	2405.5	64.3	1255	1	US-08-467-083-68
4	2405.5	64.3	1255	1	US-08-414-417B-68
5	2405.5	64.3	1255	2	US-08-486-348A-68
6	2405.5	64.3	1255	2	US-08-625-101-2
7	2405.5	64.3	1255	2	US-08-468-545B-68
8	2405.5	64.3	1255	2	US-08-356-786-2
9	2405.5	64.3	1255	3	US-08-466-680B-68
10	2405.5	64.3	1255	4	US-09-354-533-68
11	2385.5	63.8	1255	2	US-08-484-438-8
12	2224	59.5	782	2	US-09-146-283-4
13	2224	59.5	782	3	US-08-579-823A-4
14	2224	59.5	782	3	US-09-344-195-4
15	1607	43.0	419	4	US-09-630-155-2
16	1597	42.7	624	3	US-08-422-108-1
17	1597	42.7	624	4	US-08-422-734-1
18	1183	31.6	580	1	US-08-414-417B-69
19	1183	31.6	580	2	US-08-486-348A-69
20	1183	31.6	580	2	US-08-468-545B-69
21	1183	31.6	580	3	US-08-466-680B-69
22	1183	31.6	580	4	US-09-354-533-69
23	860	23.0	166	4	US-09-648-067A-1
24	761	20.4	515	2	US-09-146-283-2
25	761	20.4	515	3	US-08-579-823A-2
26	761	20.4	515	3	US-09-344-195-2
27	721	19.3	644	1	US-08-336-708A-9

28	721	19.3	911	2	US-08-484-438-10	Sequence 10, Appl
29	721	19.3	1058	2	US-08-484-438-4	Sequence 4, Appl
30	721	19.3	1210	2	US-08-484-438-7	Sequence 7, Appl
31	721	19.3	1210	2	US-08-475-035-4	Sequence 4, Appl
32	721	19.3	1308	2	US-08-484-438-2	Sequence 2, Appl
33	704.5	18.8	478	4	US-09-570-454-2	Sequence 2, Appl
34	704.5	18.8	478	4	US-09-867-521-2	Sequence 2, Appl
35	698.5	18.7	1343	6	5183884-4	Patent No. 5183884
36	697	18.6	1342	1	US-07-978-895-4	Sequence 4, Appl
37	697	18.6	1342	2	US-08-484-438-9	Sequence 9, Appl
38	697	18.6	1342	2	US-08-473-119-4	Sequence 4, Appl
39	697	18.6	1342	2	US-08-475-352-4	Sequence 4, Appl
40	697	18.6	1342	2	US-09-170-699-4	Sequence 4, Appl
41	693	18.5	274	3	US-08-469-318-144	Sequence 144, App
42	693	18.5	274	3	US-08-468-609A-144	Sequence 144, App
43	693	18.5	274	4	US-08-446-872A-144	Sequence 144, App
44	693	18.5	274	4	US-08-762-227A-144	Sequence 144, App
45	693	18.5	274	5	PCT-US95-01185-144	Sequence 144, App

ALIGNMENTS

RESULT 1

US-09-527-487-2
; Sequence 2, Application US/09527487
; Patent No. 6528060
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles
; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES
; FILE REFERENCE: 126881309200
; CURRENT APPLICATION NUMBER: US/09/527,487
; CURRENT FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-527-487-2

Query Match 64.5%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 3.7e-179;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY	35	LARGAASQVCTGDMKRLRLPASPEHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQ	94
DB	16	LPPGAASQVCTGDMKRLRLPASPEHLDMLRHLYQGCVVQGNLELYLPTNASLSFLQ	75
QY	95	DIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPG	154
DB	76	DIQEVQGYVLIAHNQVRQVPLQRLRIVRGTLFEDNYALAVLDNGDPLNNTTPTVTGASPG	135
QY	155	GLRELQLRSLTEILKGGVLIQNPQLCQVDITLWKDIFHKNNQLALTILIDNRSRACHPC	214
DB	136	GLRELQLRSLTEILKGGVLIQNPQLCQVDITLWKDIFHKNNQLALTILIDNRSRACHPC	195
QY	215	SPMKGRSCWGESSEDCQSLRTVCAGGCACRCKPLPTDCCHCCCAAGCTGPKHSDCLAC	274
DB	196	SPMKGRSCWGESSEDCQSLRTVCAGGCACRCKPLPTDCCHCCCAAGCTGPKHSDCLAC	255
QY	275	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGS----	329
DB	256	LHFHNSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC	315
QY	330	-----	329
DB	316	PLHQVETAEADGTQRCCKSPCARVCYGLGMEHLREVRVTSANIOEFGACKIFGSLA	375
QY	330	-----	329
DB	376	FLPESFDGDPASNTAPLOPEQLQVFETLEETGYLYISAWPDSLPDLVSFQNLQVIRORI	435

```
Qy 330 ----- 329
Db 436 LHNGAYSLTLOGLISWGLRSLRSLGSLALIHNNTHLCFVHTVPWDQFRNPHQALLH 495
Qy 330 ----- 329
Db 496 TANRPEDECVGEGLAGHQLCARGHCWGPQTQCVNCSQFLRGQECVEBCRVLOGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVCAHYKDPFCVACRPSGVKPDLSYMPIWK 615
Qy 330 ----- 329
Db 616 PPDEGACQPCPNCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVVLGVVFGILLI 675
Qy 330 ----- 329
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFGTIV 735
Qy 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTSPKANKELDEAYVMAGVSPYVSRLLGICLTSTVQ 795
Qy 330 ----- 329
Db 796 LVTQLMPYGCLLDHVRENRLGSGDQLLNWCQIAKMGSYLEDVRLVHRDLAARNVLVKS 855
Qy 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGKVKPIKWMVALESILRRRTHQSDVMSYGVTVWEL 915
Qy 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPPICITDVYIMVCKWIMIDSECRPRFRELVE 975
Qy 330 ----- 329
Db 976 FSRWARDQRFVVIQNEDLGPASPLDSTFYRSLLDDDDMGDLVDAEYLVYVQQGFPCDP 1035
Qy 330 --GAGGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSLAPSGAGSDVPDGLGMGAA 387
Db 1036 APGAGGMVHRRSSSTRSGGDLTLGLEPSEEBAPRSLAPSGAGSDVPDGLGMGAA 1095
Qy 388 KGLQSLPHDPSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVBPQPSREG 447
Db 1096 KGLQSLPHDPSPLQRYSEDTVPLPSETDGYVAPLTCSPQPEYVQPDVBPQPSREG 1155
Qy 448 PLPAARPAATLERAKTILSPKNGVVKDVFAFGAVENPEYLTPOGGAAPOHPHPPAFSP 507
Db 1156 PLPAARPAATLERAKTILSPKNGVVKDVFAFGAVENPEYLTPOGGAAPOHPHPPAFSP 1215
Qy 508 AFDNLYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVP 546
Db 1216 AFDNLYWDQDPPERGAAPPSTFKGTPTAENPEYLGLDVP 1254
```

RESULT 2

```
US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811.115
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match      64.5%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 3.7e-179;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

Qy 35 LARGAASQTQCTGDMKRLPASPETHLDMLRHLHYQGCQVVOGNNLELTYPNLSLSFLQ 94
Db 16 LPPGAASQTQCTGDMKRLPASPETHLDMLRHLHYQGCQVVOGNNLELTYPNLSLSFLQ 75
Qy 95 DIQSVQGVLIANQVROVPLQRLRIVRGTLQEDNEDNALAVLDNGDPLNNTPTVTGASPG 154
Db 76 DIQSVQGVLIANQVROVPLQRLRIVRGTLQEDNEDNALAVLDNGDPLNNTPTVTGASPG 135
Qy 155 GLRELQRLSLTEILKGGVLIQORNPQLCYQDTILWKDIFHKNNQLALTLIDTNRSRACHPC 214
Db 136 GLRELQRLSLTEILKGGVLIQORNPQLCYQDTILWKDIFHKNNQLALTLIDTNRSRACHPC 195
Qy 215 SPMCKGSRGWESSEDCQSLTRTVACGACRCKGPLPTDCHEQCAAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRGWESSEDCQSLTRTVACGACRCKGPLPTDCHEQCAAGCTGPKHSDCLAC 255
Qy 275 LHFNHSGICEIHCPLALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGS --- 329
Db 256 LHFNHSGICEIHCPLALVTYNTDTPESMPNPEGRYTFGASCVTACPYNYLSTDVGSCTLVC 315
Qy 330 ----- 329
Db 316 PLHNQVTAEDGTQRCCKPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLA 375
Qy 330 ----- 329
Db 376 FLPEFDDGPASNTAPLQPEQLQVFETLEEITGYLIISAWPDSLPLDLSVFQNLQVIRGI 435
Qy 330 ----- 329
Db 436 LHNGAYSLTLOGLISWGLRSLRSLGSLALIHNNTHLCFVHTVPWDQFRNPHQALLH 495
Qy 330 ----- 329
Db 496 TANRPEDECVGEGLAGHQLCARGHCWGPQTQCVNCSQFLRGQECVEBCRVLOGLPREYV 555
Qy 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVCAHYKDPFCVACRPSGVKPDLSYMPIWK 615
Qy 330 ----- 329
Db 616 PPDEGACQPCPNCTHSCVDLDDKGCAPAEORASPLTSIVSAVVGILLVVVLGVVFGILLI 675
Qy 330 ----- 329
Db 676 KRRQKIRKYTMRRLLQETELVEPLTPSGAMPNQAMRILKETELRKVKVLGSGAFGTIV 735
Qy 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTSPKANKELDEAYVMAGVSPYVSRLLGICLTSTVQ 795
Qy 330 ----- 329
Db 796 LVTQLMPYGCLLDHVRENRLGSGDQLLNWCQIAKMGSYLEDVRLVHRDLAARNVLVKS 855
Qy 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGKVKPIKWMVALESILRRRTHQSDVMSYGVTVWEL 915
Qy 330 ----- 329
Db 916 MTFGAKPYDGIPIAREIPDLLEKGERLPQPPICITDVYIMVCKWIMIDSECRPRFRELVE 975
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 75.9936 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: us-09-821-883-2

Perfect score: 3739

Sequence: 1 MRAAPLLARAAASLSGLF.....EPVQEGAPPPAAHHHHH 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3739	100.0	690	4	AAE13109 Human HER
2	3473.5	92.9	697	4	AAE13111 Human HER
3	2972	79.5	555	4	AAE13108 Human HER
4	2957.5	79.1	564	4	AAE13110 Human HER
5	2573.5	68.8	919	3	AAE21203 Human HER
6	2573.5	68.8	919	5	AAE13114 Human HER
7	2410.5	64.5	1255	3	AAE13109 Human HER
8	2410.5	64.5	1255	4	AAE13111 Human HER
9	2410.5	64.5	1255	5	AAE13108 Human HER
10	2410.5	64.5	1255	4	AAE13110 Human HER
11	2410.5	64.5	1255	5	AAE21203 Human HER
12	2410.5	64.5	1255	5	AAE26366 Human HER
13	2410.5	64.5	1255	6	AAU74545 Human HER
14	2410.5	64.5	1255	6	ABP74708 Human HER
15	2410.5	64.5	1255	6	AAE38390 Human HER
16	2410.5	64.5	1255	6	ADA38143 Human HER
17	2410.5	64.5	1255	7	ADA37255 Human HER
18	2410.5	64.5	1255	7	ADB67621 Human HER
19	2405.5	64.3	1255	2	AAW01111 Human HER
20	2405.5	64.3	1255	2	AAW92406 Human HER
21	2405.5	64.3	1255	3	AAV84780 Human HER
22	2405.5	64.3	1255	3	AAE21198 Human HER
23	2405.5	64.3	1255	4	AAE88267 Human HER
24	2405.5	64.3	1255	4	AAE85458 Human HER
25	2405.5	64.3	1255	5	AAE20479 Human HER

26	2405.5	64.3	1255	5	AAU77114 Human HER
27	2405.5	64.3	1255	5	AAE51143 Human HER
28	2405.5	64.3	1255	5	AAE24067 Human HER
29	2405.5	64.3	1255	6	ABR43687 Human C-e
30	2405.5	64.3	1255	7	ABR82066 Human HER
31	2405.5	64.3	1255	7	ADC09593 Human HER
32	2405.5	64.3	1255	7	ADD25484 Binding d
33	2405.5	64.3	1255	7	ADG63281 Human PRO
34	2405.5	64.3	1255	7	ADG76190 Human HER
35	2400.5	64.2	1433	2	AAE39568 Sequence
36	2398.5	64.1	1253	7	ADC35106 Human bre
37	2334.5	62.4	1223	5	AAU98923 Human bre
38	2224	59.5	782	2	AAW19764 Her2-GM-C
39	2188.5	58.5	479	4	AAE13112 Human HER
40	2130	57.0	926	5	AAE51153 Mouse HER
41	2129	56.9	920	5	AAE51152 Mouse HER
42	2092.5	56.0	1200	3	AAE21208 Human HER
43	1961	52.4	1256	3	AAE21206 Mouse HER
44	1961	52.4	1256	4	AAE62860 Amino aci
45	1961	52.4	1256	5	AAE51151 Mouse HER

ALIGNMENTS

RESULT 1

AAE13109
ID AAE13109 standard; protein; 690 AA.

XX AAE13109;

DT 28-JAN-2002 (first entry)

DE Human HER500-hGM-CSF fusion protein construct.

KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW HER500-hGM-CSF fusion protein.

XX Homo sapiens.

OS Synthetic.

OS Chimeric.

PN WO200174855-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010515.

XX 30-MAR-2000; 2000US-0193504P.

XX (DEND-) DENDREON CORP.

XX Laus R, Vidovic D, Graddis T;

XX WPI; 2001-662965/76.

XX N-PSDB; AAD21565.

XX An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

XX Claim 7; Page 26; 59pp; English.

XX The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular

CC immune response to the antigen. IFP or superactivated dendritic cells are
CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
CC associated with a particularly antigen. The present sequence is HER500
CC rGM-CSF fusion protein construct which comprises human PAP signal
CC sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
CC sequence, mature HER-2 membrane distal extracellular and intracellular
CC domains, an Ala Ala linker, a mature human granulocyte- macrophage colony
CC stimulating factor (GM-CSF) sequence and a C-terminal tag
XX
XX Sequence 690 AA;

Query Match 100.0%; Score 3739; DB 4; Length 690;
Best Local Similarity 100.0%; Pred. No. 7.7e-224;
Matches 690; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRAAPLLARAAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGDMKRLRPASPET 60
Db 1 MRAAPLLARAAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGDMKRLRPASPET 60
Qy 61 HLDMLRHLVYQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLVYQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
Qy 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Qy 241 GGCARCKGLPTDCCHEOCACGCTGPKHSDCLACLFHNSHIGELHCPALVTYNTDFES 300
Db 241 GGCARCKGLPTDCCHEOCACGCTGPKHSDCLACLFHNSHIGELHCPALVTYNTDFES 300
Qy 301 MPNPEGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRSSSTRSGGDLTLGLEPSEE 360
Db 301 MPNPEGRYTFGASCVTACPYNYLSTDVSGAGGVMVHRRSSSTRSGGDLTLGLEPSEE 360
Qy 361 EAPRSPLAPSGAGSDVFDGLGMAAGKLSLPHDPSPLQRYSEDPTVPLPSETDGYV 420
Db 361 EAPRSPLAPSGAGSDVFDGLGMAAGKLSLPHDPSPLQRYSEDPTVPLPSETDGYV 420
Qy 421 APLTCSPOEYVNDVDPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVDVFAFG 480
Db 421 APLTCSPOEYVNDVDPQPPSPREGPLPAARPAAGATLERAKTILSPKNGVVDVFAFG 480
Qy 481 GAVENPEYLTQGGAAPOPHPPAFSPAFDNLVYWDQDPPRGAPPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTQGGAAPOPHPPAFSPAFDNLVYWDQDPPRGAPPSTFKGTPTAENPEY 540
Qy 541 LGLDVPAAPARSPPSTQPEHVNAIQEARLLNLSRDTAEMNETVEVISEMFDLQEP 600
Db 541 LGLDVPAAPARSPPSTQPEHVNAIQEARLLNLSRDTAEMNETVEVISEMFDLQEP 600
Qy 601 TCLQTRLELYKQGLRGSITLKGPLTMMASHYKQHCPTPTSCATQIITPESKFNLD 660
Db 601 TCLQTRLELYKQGLRGSITLKGPLTMMASHYKQHCPTPTSCATQIITPESKFNLD 660
Qy 661 FLLVTFPCWEPVQEGAPPPAPAAHHHHH 690
Db 661 FLLVTFPCWEPVQEGAPPPAPAAHHHHH 690

RESULT 2
AAE13111
ID AAE13111 standard; protein; 697 AA.

XX AC AAE13111;
XX 28-JAN-2002 (first entry)
XX DE Human HER500-rGM-CSF fusion construct comprising OVA-derived peptide.

XX Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human; GM-CSF;
KW HER-2 protein; granulocyte-macrophage colony stimulating factor;
KW ovalbumin-derived octapeptide; OVA; rat; HER500-rGM-CSF fusion protein.
XX
OS Homo sapiens.
OS Rattus norvegicus.
OS Unidentified.
OS Chimeric.
PN WO200174855-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US010515.
XX 30-MAR-2000; 2000US-0193504P.
XX (DEND-) DENDREON CORP.
XX Laus R, Vidovic D, Graddis T;
PI WPI; 2001-662965/76.
XX N-PSDB; AAD21567.
XX An immunostimulatory fusion protein comprising the intracellular domain
of HER-2 and an antigen elicits an immune response to the antigen and is
useful for the treatment of associated cancer associated.
XX Claim 7; Page 27; 59pp; English.
XX The invention relates to immunostimulatory fusion proteins (IFP) and
nucleic acid molecules encoding such proteins. The IFPs comprise a
polypeptide antigen component and an immunostimulatory component derived
from the intracellular domain of HER-2 protein which is effective to
elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
immune response to the antigen. IFP or superactivated dendritic cells are
used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
associated with a particularly antigen. The present sequence is HER500
rGM-CSF fusion protein construct which comprises human PAP signal
sequence, mature PAP protein, an Ala Arg linker, human HER-2 signal
sequence, mature HER-2 membrane distal extracellular domain, an Ala
linker, an ovalbumin (OVA)-derived immunodominant octapeptide, HER-2
membrane distal intracellular domain, an Ala Ala linker, a mature rat
granulocyte-macrophage colony stimulating factor (GM-CSF) sequence and a
C-terminal tag
XX
SQ Sequence 697 AA;

Query Match 92.9%; Score 3473.5; DB 4; Length 697;
Best Local Similarity 91.7%; Pred. No. 2.4e-207;
Matches 641; Conservative 19; Mismatches 28; Indels 11; Gaps 2;
Qy 1 MRAAPLLARAAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGDMKRLRPASPET 60
Db 1 MRAAPLLARAAASLSGLFLFLFWLDRSVLAKELARGAASQVCTGDMKRLRPASPET 60
Qy 61 HLDMLRHLVYQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
Db 61 HLDMLRHLVYQGCQVQVQGNLELYLPTNASLSFLQDIOEVQGVLIHNVQVPLQRLRI 120
Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQLRSLEILKGGVLIQRLNPOL 180
Qy 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDTNRSRACHPCSPMCKGRCWGSESDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:57:41 ; Search time 3742.37 Seconds
(without alignments)
4428.608 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLARAASLSGLFLF.....ENPEYGLDVPAAAHHHH 555

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0982183/runat_09092004_105127_6643/app_query.fasta_1.3100
-DB=EST -QFMT=fastcap -SUFFIX=rest -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -List=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0982183 @CGN 1 1 11571 @runat_09092004_105127_6643 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=110 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_phg:*
27: em_ges_vrl:*
28: gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1964	65.0	4323	11	AK031099	AK031099 Mus muscu
2	1947.5	64.5	4463	11	AK083669	AK083669 Mus muscu
3	1288	42.6	1016	12	BM562913	BM562913 AGENCOURT
4	1175	38.9	1129	12	BM802792	BM802792 AGENCOURT
5	1163	38.5	871	14	CA488274	CA488274 AGENCOURT
6	1163	38.5	885	14	CA455074	CA455074 AGENCOURT
7	1156	38.3	4715	11	AP318349	AP318349 Homo sapi
8	1137.5	37.7	893	14	CA455141	CA455141 AGENCOURT
9	1090	36.1	1004	14	CD515356	CD515356 AGENCOURT
10	1087	36.0	894	14	CA454131	CA454131 AGENCOURT
11	1080	35.8	653	12	BM721340	BM721340 UI-E-E01-
12	1073	35.5	588	13	BX478931	BX478931 DKFZp686A
13	1072	35.5	902	14	CA488868	CA488868 AGENCOURT
14	1046	34.6	569	9	AL701765	AL701765 DKFZp686P
15	1027	34.0	583	13	EX479114	EX479114 DKFZp686P
16	1016.5	33.7	795	14	CA328613	CA328613 UI-M-PY0-
17	1012.5	33.5	3110	11	AK031542	AK031542 Mus muscu
18	1004	33.2	808	14	CA489534	CA489534 AGENCOURT
19	997	33.0	616	14	CD721801	CD721801 oJ01f07.Y
20	976	32.3	943	13	BQ958632	BQ958632 AGENCOURT
21	976	32.3	1988	11	BC023725	BC023725 Mus muscu
22	976	32.3	3372	11	BC046553	BC046553 Mus muscu
23	931	30.8	813	12	BI557797	BI557797 603236977
24	926	30.7	855	13	BU594980	BU594980 AGENCOURT
25	915.5	30.3	852	14	CD516283	CD516283 AGENCOURT
26	915	30.3	614	10	AW370693	AW370693 QV1-BT026
27	912	30.2	791	12	BI154872	BI154872 602902857
28	910	30.1	717	14	CB598701	CB598701 AGENCOURT
29	901	29.8	491	12	BM790293	BM790293 K-EST0070
30	886	29.3	906	14	CA454570	CA454570 AGENCOURT
31	883	29.2	998	12	BI649877	BI649877 603296516
32	881	29.2	685	9	AU123871	AU123871 AU123871
33	879.5	29.1	932	14	CA487981	CA487981 AGENCOURT
34	868	28.7	609	10	AW701942	AW701942 ug93b02.Y
35	862	28.5	621	12	BG283493	BG283493 602407782
36	855	28.3	786	12	BI155788	BI155788 602904360
37	846	28.0	887	13	BQ717097	BQ717097 AGENCOURT
38	835.5	27.7	649	9	AA496412	AA496412 zv37a02.f
39	817	27.1	533	14	CB437620	CB437620 685491.MA
40	817	27.1	552	12	BG277542	BG277542 ux45d07.Y
41	811	26.9	670	9	AI906012	AI906012 RC-BT105-
42	808	26.8	876	13	BQ769889	BQ769889 UI-M-P10-
43	803	26.6	887	14	CA980253	CA980253 AGENCOURT
44	802	26.6	514	10	BF998814	BF998814 QV0-GN014
45	797.5	26.4	678	12	BI555157	BI555157 603236396

ALIGNMENTS

RESULT 1
AK031099

LOCUS
DEFINITION

AK031099 4323 bp mRNA linear HTC 18-SEP-2003
Mus musculus 13 days embryo forelimb cDNA, RIKEN full-length
enriched library, clone:530404N10 product:v-erb-b2 erythroblastic
leukemia viral oncogene homolog 2, neuro/glioblastoma derived
oncogene homolog (avian), full insert sequence.

ACCESSION
AK031099

VERSION
AK031099.1 GI:26082143

KEYWORDS
HTC; CAP trapper.

SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayaishizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayaishizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, I., Ozawa, K., Tanaka, I., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayaishizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 585-590 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4323)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayaishida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayaishizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayaishizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers
1. .4323
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

/db_xref="FANTOM DB:5930404N10"
/db_xref="MGI:2394277"
/db_xref="taxon:10090"
/clone="5930404N10"
/tissue_type="forelimb"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
48. .3818
/note="putative
v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)
(MGI:95410, GB|U71126, evidence: BLASTN, 99%, match=449)"

misc_feature
48. .3818

ORIGIN
Alignment Scores:
Pred. No.: 1.25e-112 Length: 4323
Score: 1964.00 Matches: 437
Percent Similarity: 36.17% Conservative: 26
Best Local Similarity: 34.14% Mismatches: 77
Query Match: 65.03% Indels: 740
DB: 11 Gaps: 3

US-09-821-883-1 (1-555) x AK031099 (1-4323)

Qy 3 AlaAlaProLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPheLeuLeu 22
||||| :
Db 33 GCCGAGTATCATCATGGAGCTGGCGCTGGTGGGGGTTCCTCTCGCCCTC 92
||||| :
Qy 23 PhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaArgGlyAlaAlaSerThr 42
||||| :
Db 93 -----CTGTCCTCCCGGAGCGCGGTACC 116
||||| :
Qy 43 GlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThrHisLeu 62
||||| :
Db 117 CAAAGTGTGTACCGGTACCGATGAAGTTGGACTCTCTGCTGCTGAGACCCACTG 176
||||| :
Qy 63 AspMetLeuArgHisLeuTyrglnGlyCysglnValValGlnGlnGlyAsnLeuGluLeuThr 82
||||| :
Db 177 GACATGTTGCGCCACCTCTACCGAGGTGTGAGTGTGTCAGGGCAATTTGGAGCTTACC 236
||||| :
Qy 83 TyrlleuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnGluValGlnGlyTy 102
||||| :
Db 237 TACCTGCGCCCAATGCGAGCTCTCATCTCTGTCAGGACATCCAGGAAGTCCAGGATAC 296
||||| :
Qy 103 ValIleleAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgGlyIleValArg 122
||||| :
Db 297 ATGCTCATCTGCTCACACCGAGTGAACACGTCCTCCACTGTCAGAGGTTCGCGATCGTGA 356
||||| :
Qy 123 GlyThrGlnLeuPheGluAspAsnTyrlleAlaAlaValLeuAspAsnGlyAspProLeu 142
||||| :
Db 357 GGGACTCAGCTCTTTGAGGACAGTATGCCCTGGCTGTGTAGACACCGAGACCTTTG 416
||||| :
Qy 143 AsnAsn---ThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGlnLeu 161
||||| :
Db 417 GACACGCTCACCGCGCCCGCCAGGAGAGAACCCAGAGGGCTGCGGAGCTGCAGCTT 476
||||| :
Qy 162 ArgSerLeuThrGluIleLeuLysGlyValleuIleGlnArgAsnProGlnLeuCys 181
||||| :
Db 477 CGAAGTCTCAGAGATCTTTGAAGGGAGGAGTTTGTATCGGTGGGAACCTTCAGCTCTGC 536
||||| :
Qy 182 TyrlleuAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeuThr 201
||||| :
Db 537 TACCGAGACATGTTTGTGGAGGATGTCTCCGTAAAGATTAACAGCTGGCTCTCTGTC 596
||||| :
Qy 202 LeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGlySer 221
||||| :
Db 597 GACATGGACACCAATCGTTCGGGGCTGTCCACCTTGTGTGCCCCCAACCTGCAAGACAAT 656
||||| :
Qy 222 ArgCysTrpGlyGluSerSerGluAspCysGlnSerLeuThrValCysAlaGly 241
||||| :
Db 657 CACTGTTGGGGGTGAGAGTCTCTGAAGACTGTGAGATCTTGACTGGCACCACTCTGACTAGT 716
||||| :

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 15:10:16 ; Search time 671.049 Seconds
(without alignments)
4161.385 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSGLFLF.....ENPEYGLDVPAHHHHH 555

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3304383 seqs, 2515761380 residues

Total number of hits satisfying chosen parameters: 6608766

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09821883/runat_09092004_105129_6736/app_query.fasta_1.3100
-DB=Published Applications NA_QFMT=fastap -SUFFIX=rnppb -MINMATCH=0.1
-LOOFCPU=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=5 -MODE=LOCAL -OUTFMT=ptc -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -ALIGN=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09821883@cgn 1.1949 @runat_09092004_105129_6736
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT_DSBELOC=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRSADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
------------	-------	--------------	----	-------------

ALIGNMENTS

RESULT 1

US-09-821-883-6
; Sequence 6, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; FILE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821.883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct

1	3020	100.0	1665	9	US-09-821-883-6	Sequence 6, Appl
2	3005.5	99.5	1692	9	US-09-821-883-8	Sequence 8, Appl
3	2972	98.4	2070	9	US-09-821-883-7	Sequence 7, Appl
4	2957.5	97.9	2091	9	US-09-821-883-9	Sequence 9, Appl
5	2419	80.1	4543	9	US-09-769-508-1	Sequence 1, Appl
6	2416	80.0	4530	9	US-09-877-177-11	Sequence 11, Appl
7	2416	80.0	4530	13	US-10-392-113-45	Sequence 45, Appl
8	2416	80.0	4530	13	US-10-426-836-11	Sequence 11, Appl
9	2416	80.0	4530	15	US-10-177-293-125	Sequence 125, Appl
10	2416	80.0	4530	15	US-10-007-926A-119	Sequence 119, Appl
11	2416	80.0	4530	15	US-10-338-730-1	Sequence 1, Appl
12	2416	80.0	4530	15	US-10-101-510-124	Sequence 124, Appl
13	2416	80.0	4530	16	US-10-116-275-131	Sequence 27, Appl
14	2416	80.0	4530	16	US-10-272-437A-27	Sequence 27, Appl
15	2416	80.0	4530	16	US-10-117-937-595	Sequence 595, Appl
16	2416	80.0	4530	16	US-10-159-563-208	Sequence 208, Appl
17	2416	80.0	4530	16	US-10-435-696-10	Sequence 10, Appl
18	2416	80.0	4530	17	US-10-734-564-59	Sequence 59, Appl
19	2416	80.0	4642	15	US-10-198-846-10896	Sequence 10896, A
20	2413.5	79.9	9274	9	US-09-811-123-7	Sequence 7, Appl
21	2413.5	79.9	9274	9	US-09-811-115-1	Sequence 1, Appl
22	2411	79.8	4473	10	US-09-441-411-5	Sequence 5, Appl
23	2411	79.8	4473	15	US-10-146-473-32	Sequence 32, Appl
24	2411	79.8	4473	15	US-10-207-655-44	Sequence 44, Appl
25	2411	79.8	4473	15	US-10-101-510-81	Sequence 81, Appl
26	2410.5	79.8	3765	15	US-10-207-498-5	Sequence 5, Appl
27	2410.5	79.8	3768	9	US-09-811-123-8	Sequence 8, Appl
28	2410.5	79.8	3768	9	US-09-811-115-2	Sequence 2, Appl
29	2410.5	79.8	3768	13	US-09-984-092-3	Sequence 3, Appl
30	2410.5	79.8	3768	13	US-10-280-576-3	Sequence 3, Appl
31	2410.5	79.8	3768	17	US-10-441-779C-3	Sequence 3, Appl
32	2405.5	79.7	3768	9	US-09-854-356-9	Sequence 9, Appl
33	2405.5	79.7	3768	9	US-09-930-125-1	Sequence 1, Appl
34	2405.5	79.7	3768	13	US-09-765-973-1	Sequence 1, Appl
35	2405.5	79.7	3768	15	US-10-313-644-1	Sequence 1, Appl
36	2386	79.0	4606	10	US-09-971-392-70	Sequence 70, Appl
37	1961	64.9	3771	9	US-09-854-356-11	Sequence 11, Appl
38	1956.5	64.8	3955	9	US-09-870-759-117	Sequence 117, Appl
39	1956.5	64.8	3955	9	US-09-854-356-10	Sequence 10, Appl
40	1956.5	64.8	3955	10	US-09-751-708A-117	Sequence 117, Appl
41	1795	59.4	1437	9	US-09-821-883-10	Sequence 10, Appl
42	1641.5	54.4	2132	16	US-10-412-804A-3	Sequence 3, Appl
43	1639	54.3	2149	16	US-10-412-804A-9	Sequence 9, Appl
44	1636	54.2	2164	16	US-10-412-804A-5	Sequence 5, Appl
45	1587	52.5	867	9	US-09-821-883-24	Sequence 24, Appl

US-09-821-883-6

Alignment Scores:

Pred. No.:	4,938-272	Length:	1655
Score:	3020.00	Matches:	555
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-821-883-1 (1-555) x US-09-821-883-6 (1-1665)

Qy	1	Met	Arg	Ala	Ala	Pro	Leu	Leu	Leu	Ala	Arg	Ala	Ala	Ser	Leu	Ser	Leu	Gly	Phe	Leu	Phe	20													
Db	1	ATG	AGAG	CTG	CAC	CCC	TCT	CCT	TG	CC	AGG	GAG	CA	AGC	CTT	TAG	CGT	TTC	TGG	CTT	CTG	TTT	60												
Qy	21	Leu	Leu	Phe	Phe	Thr	Leu	Asp	Arg	Ser	Val	Leu	Ala	Leu	Val	Ser	Leu	Ala	Arg	Gly	Ala	Ala	40												
Db	61	CTG	CTT	TTT	TTT	TCT	GG	CT	TAG	CCG	AAG	TGT	TAC	TGC	CA	AGG	AGT	TGG	CG	CGG	CGG	CGG	CGG	120											
Qy	41	Ser	Thr	Gln	Val	Cys	Thr	Gly	Thr	Asp	Met	Leu	Ser	Leu	Arg	Leu	Pro	Ala	Ser	Pro	Glu	Thr	60												
Db	121	TCG	AC	CC	CA	AG	TGT	CAC	CGG	CAC	AG	CAC	ATG	AG	CT	TGC	GC	TCT	CC	CT	CC	AG	TCC	180											
Qy	61	His	Leu	Asp	Met	Leu	Arg	His	Leu	Tyr	Gln	Gly	Cys	Gln	Val	Val	Gln	Gly	Asn	Leu	Glu	240													
Db	181	CAC	CT	TG	CA	CA	TG	CT	CG	CAC	CT	TAC	CAG	GCT	TGC	CAG	TGT	G	TG	CAG	GAA	CTT	GGA	240											
Qy	81	Leu	Thr	Tyr	Leu	Pro	Thr	Asn	Ala	Ser	Leu	Ser	Phe	Leu	Gln	Asp	Leu	Gln	Val	Gln	100														
Db	241	CTC	AC	CT	TAC	CT	TGC	CA	CC	AAT	TGC	CA	GC	CT	TGT	CT	TGC	AG	AT	TCC	AG	AG	TTC	300											
Qy	101	Gly	Tyr	Val	Leu	Leu	Ala	His	Asn	Gln	Val	Arg	Gln	Val	Pro	Leu	Gln	Arg	Leu	Arg	Glu	120													
Db	301	GGC	TAC	TG	CT	CAT	CG	CT	CAC	CA	CC	AG	TG	AG	CAG	CT	CC	CA	CT	CAG	AGG	CTG	CG	360											
Qy	121	Val	Arg	Gly	Thr	Gln	Leu	Phe	Glu	Asp	Asn	Tyr	Ala	Leu	Ala	Val	Leu	Asp	Asn	Gly	Asp	140													
Db	361	GTG	CG	AGG	CAC	CC	CAG	CT	CTT	TG	AG	GAC	AA	CT	AT	G	CC	T	G	CC	T	G	CC	420											
Qy	141	Pro	Leu	Asn	Asn	Thr	Thr	Pro	Val	Thr	Gly	Ala	Ser	Pro	Gly	Gly	Leu	Arg	Glu	Leu	Gln	160													
Db	421	CCG	CT	GAA	CA	AAT	TAC	CA	CC	CC	CT	TGT	CA	CA	G	GG	GG	CC	T	CC	CAG	AGG	CTG	CGG	480										
Qy	161	Leu	Arg	Ser	Leu	Thr	Glu	Leu	Leu	Leu	Ser	Gly	Val	Leu	Leu	Leu	Ser	Ala	Arg	Asn	Pro	Gln	Leu	180											
Db	481	CTT	TG	AG	CC	CT	CAC	AG	AT	CTT	GAA	GG	GGG	CT	T	TG	AT	CAG	CGG	AA	CCCC	CAG	CTC	540											
Qy	181	Cys	Tyr	Gln	Asp	Thr	Ile	Leu	Trp	Leu	Asp	Leu	Phe	His	Leu	Asn	Gln	Leu	Ala	Leu	200														
Db	541	TG	CT	TAC	CAG	GAC	CAG	TAT	T	TG	TGA	AGG	AG	AT	CTT	CC	CA	AG	AA	CA	AC	CAG	CTG	GC	600										
Qy	201	Thr	Leu	Ile	Asp	Thr	Asn	Arg	Ser	Arg	Ala	Cys	His	Pro	Cys	Ser	Pro	Met	Cys	Leu	Ser	220													
Db	601	AC	ACT	TG	TAG	AC	CA	CA	CC	CG	CT	CT	CG	GC	CT	CC	CA	CC	CC	CT	TGT	TCT	CCG	AT	G	T	G	T	A	A	G	G	G	C	660
Qy	221	Ser	Arg	Cys	Trp	Gly	Glu	Ser	Ser	Glu	Asp	Cys	Gln	Ser	Leu	Thr	Arg	Thr	Val	Cys	Ala	240													
Db	661	TCC	CG	CT	CT	CGG	AG	AG	ATT	CT	GAG	ATT	G	T	CAG	GC	CT	G	AG	CG	CAC	CT	G	T	C	T	G	T	G	G	C	720			
Qy	241	Gly	Gly	Cys	Ala	Arg	Cys	Leu	Ser	Gly	Pro	Leu	Pro	Thr	Asp	Cys	Cys	His	Glu	Gln	Cys	Ala	260</												

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 06:51:36 ; Search time 95.6744 Seconds
(without alignments)
3219.229 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAPALLARASLSLGFLP.....ENPEYLGLDVPAAAHHHHH 555

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n model -DEV=xlh
-Q/cgn2_1/USPTO_spool/US09821883/runat_09092004_105128_6657/app_query.fasta_1.3100
-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09821883 @CGN 1 1 215 @runat_09092004_105128_6657 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2416	80.0	4530	1	US-08-229-515A-9
2	2416	80.0	4530	1	Sequence 9, Appli
3	2416	80.0	4530	4	US-08-645-865-9
4	2416	80.0	4530	4	Sequence 4, Appli
5	2416	80.0	4530	4	US-09-527-487-1
6	2413.5	79.9	9274	4	US-09-877-177A-11
7	2411	79.8	4473	2	US-09-811-115-1
8	2411	79.8	4473	3	US-09-056-105-26
9	2410.5	79.8	4473	4	US-09-663-834A-3
10	2405.5	79.7	3768	2	US-09-811-115-2
11	2405.5	79.7	3768	2	US-08-625-101-1
12	2405.5	79.7	3768	2	US-08-356-786-1

13	1956.5	64.8	3955	1	US-08-229-515A-14
14	1956.5	64.8	3955	1	US-08-645-865-14
15	1639	54.3	2385	2	US-09-146-283-3
16	1639	54.3	2385	3	US-08-579-823A-3
17	1639	54.3	2385	3	US-09-344-195-3
18	1624	53.8	1872	3	US-08-422-108-2
19	1624	53.8	1872	4	US-08-422-734-2
20	725.5	24.0	1868	1	US-08-658-883B-1
21	725.5	24.0	1868	4	US-09-676-610B-26
22	724.5	24.0	1593	4	US-09-676-610B-25
23	724.5	24.0	5532	2	US-08-475-035-3
24	724.5	24.0	5532	4	US-09-676-610B-17
25	721	23.9	5484	3	US-09-632-580A-3
26	721	23.9	5501	1	US-08-484-438-1
27	721	23.9	5555	1	US-08-484-438-3
28	714.5	23.7	4905	1	US-07-978-895-3
29	714.5	23.7	4905	1	US-08-473-119-3
30	714.5	23.7	4905	2	US-08-475-352-3
31	714.5	23.7	4905	4	US-09-170-699-3
32	714.5	23.7	4975	3	US-09-630-706-3
33	705.5	23.4	1958	4	US-09-570-454-1
34	705.5	23.4	1958	4	US-09-867-521-1
35	623	20.6	4545	6	5183884-3
36	493	16.3	322	1	US-08-421-356-1
37	493	16.3	322	4	US-09-046-783-1
38	260.5	8.6	4149	2	US-08-737-715-1
39	260.5	8.6	4723	4	US-09-023-655-1137
40	257	8.5	2533	3	US-09-149-922-5
41	252.5	8.4	4975	2	US-08-249-687C-1
42	252.5	8.4	4989	2	US-08-666-392A-3
43	252.5	8.4	4989	2	US-08-625-819-1
44	252.5	8.4	4989	3	US-08-755-558-4
45	252.5	8.4	4989	3	US-08-746-559A-1

ALIGNMENTS

RESULT 1
US-08-229-515A-9
; Sequence 9, Application US/08229515A
; Patent No. 551885
; GENERAL INFORMATION:
; APPLICANT: RAZIUDIN
; APPLICANT: SARKAR, FAZLUL H
; TITLE OF INVENTION: ERBB2 PROMOTER BINDING PROTEIN IN
; TITLE OF INVENTION: NEOPLASTIC DISEASE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG PC
; STREET: 127 Peachtree Street, Suite 1200
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,515A
; FILING DATE: 19 APR 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRYMAN, DAVID G
; REGISTRATION NUMBER: 33,438
; REFERENCE/DOCKET NUMBER: 1414.608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-688-0770
; TELEFAX: 404-688-9880
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4530 base pairs

;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-229-515A-9

Alignment Scores:

Pred. No.: 1,67e-178 Length: 4530
Score: 2416.00 Matches: 520
Percent Similarity: 41.14% Conservative: 5
Best Local Similarity: 40.75% Mismatches: 17
Query Match: 80.00% Indels: 734
DB: 1 Gaps: 3

US-09-821-883-1 (1-555) x US-08-229-515A-9 (1-4530)

```
Qy 2 ArgAlaAlaProLeuLeuLeuAlaArgAla-----AlaSerLeuSerLeuGly 17
Db 94 CGCGCCCTCCAGCCGGGTCAGCCGGAGCATGGGGCCGGAGCGCGAGTGAGCACCATG 153
Qy 18 PheLeuPheLeuLeuPhePheTyrLeuAspArgSerValLeuAlaLysGluLeuAlaArg 37
Db 154 GAGCTGGCGGCTTGTGCCGCTG-----GGGCTCCTCTCGGCCCTTGTCCCCC 204
Qy 38 GlyAlaAlaSerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSer 57
Db 205 GGAGCCCGAGACCCCAAGTGTGCACCGGCACAGACATGAAGCTGGCGCTCCCTGCCAGT 264
Qy 58 ProGluThrHisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGly 77
Db 265 CCGGAGACCCACCTGGACATGCTCCGCCACCTCTTACCAGGGCTGCCAGGTGGTGCAGGA 324
Qy 78 AsnLeuGluLeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGln 97
Db 325 AACCTGGAATCACTACCTACCTGCCCCACCAATGCCAGCCTGTCTTCTGCAAGGATATCCAG 384
Qy 98 GluValGlnGlyTyrValLeuIleAlaHisAsnGlnValArgGlnValProLeuGlnArg 117
Db 385 GAGGTGCAGGGCTACGTGCTCATCGCTCACACCAAGTAGGAGGAGTCCCACCTGCAGAGG 444
Qy 118 LeuArgIleValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAsp 137
Db 445 CTGCGGATTGTGGAGGACCCAGCTCTTTGAGGACAACTATGCTTGGCGCTGGCGGTAGAC 504
Qy 138 AsnGlyAspProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArg 157
Db 505 AATGAGACCCGCTGAACAATACACCCCTGTACAGGGGCTCCCGAGGAGGCTTGC 564
Qy 158 GluLeuGlnLeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsn 177
Db 565 GAGCTGCAGCTTCGAGCCCTCACAGATCTTGAAAGGAGGGGTCTTGATCCAGCGGAAC 624
Qy 178 ProGlnLeuCysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnAsnGln 197
Db 625 CCCAGCTCTGCTACCAAGGACAGCATTTGTGGAAGGACATCTTCCACAAGAAACAACAG 684
Qy 198 LeuAlaLeuThrIleLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMet 217
Db 685 CTGGCTCTCACTGATAGACACCAACCGCTCTCGGGCCCTGCCACCCCTGTCTCCGATG 744
Qy 218 CysLysGlySerArgCysTyrGlyGluSerSerGluAspCysGlnSerLeuThrArgThr 237
Db 745 TGTAAAGGCTCCCGCTGCTGGGAGAGAGTCTTGAGGATTTGTAGAGCCTGACGGGACT 804
Qy 238 ValCysalaGlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGlu 257
Db 805 GTCTGTCCGGTGGCTGTGCCCGCTGCAGGGGCCACTGCCCACTGACTGCTGCCATGAG 864
Qy 258 GlnCysAlaIaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPhe 277
Db 865 CAGTGTGCTCCGGCTGCACGGGGCCCCAAGACCTCTGACTGCTGCTGGCTGCCCTCCACTTC 924
Qy 278 AsnHisSerGlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThr 297
```

```
Db 925 AACCAAGAGTGCATCTGTGAGCTGCACATGCCACGCCCTGTGTACACACACAGACACG 984
Qy 298 PheGluSerMetProAsnProGluGlyArgTyrThrPheGlyValaSerCysValThrAla 317
Db 985 TTTGAGTCCATGCCCCAATCCGAGGCGCGGTATACATTCTGGCGCCAGCTGTGACTGCC 1044
Qy 318 CysProTyrAsnTyrLeuSerThrAspValGlySer----- 329
Db 1045 TGTCCCTACAACTACCTTTCTACGGAGCTGGGATCTGTGACCCCTCGTCTGCCCTGCAC 1104
Qy 329 ----- 329
Db 1105 AACCAAGAGTGCACAGCAGAGGATGGAACACACAGCGGTGTGAGAAGTGCAGAACCCCTGT 1164
Qy 329 ----- 329
Db 1165 GCCGAGTGTGCTATGGTCTGGGCATGGAGCATCTTGGAGAGTGGAGGAGTTACCACT 1224
Qy 329 ----- 329
Db 1225 GCCAATATCCAGGAGTTTGTGTGCTGCAAGAAGATCTTTGGGAGCCTTGGCATTTCTGCCG 1284
Qy 329 ----- 329
Db 1285 GAGAGCTTTGATGGGACCCAGCCTCCAACTGCCCCGCTCCAGCCAGAGAGCTCCAA 1344
Qy 329 ----- 329
Db 1345 GTTTTGGAGACTCTGGAAGAGATCAAGGTTACTATATCTCAGCATGCGCGGACAGC 1404
Qy 329 ----- 329
Db 1405 CTGCTGACCTCAGCGCTCTTCCAGAACCTGCAAGTAATCCGGGAGCGAATTTCTGCACAAT 1464
Qy 329 ----- 329
Db 1465 GGCGCCTACTCGCTGACCCCTGCAAGGGCTGGGCATCAGCTGGGTGGGGCTCGCGTCACTG 1524
Qy 329 ----- 329
Db 1525 AGGGAATCGGGCAGTGGATCGGCCCTCATCCACATAACACCACTCTGCTTCTGCTGCAC 1584
Qy 329 ----- 329
Db 1585 ACGTGCCCTGGGACCCAGCTCTTTTCGGAACCCGCACCAAGCTCTGTCTCCACACTGCCAAC 1644
Qy 329 ----- 329
Db 1645 CGGCCAGAGGACGAGTGTGTGGCGGAGGGCTTGGCCCTGCCACCAAGCTGTGCGCCCGAGGG 1704
Qy 329 ----- 329
Db 1705 CACTGTGGGGTCCAGGGGCCACCCAGTGTGTCACTGCAGCCAGTTCTTTCGGGGGCCAG 1764
Qy 329 ----- 329
Db 1765 GAGTGTGGAGGAATCCCGAGTACTGACAGGGGCTCCCCAGGGAGTATGTGAATGCCAGG 1824
Qy 329 ----- 329
Db 1825 CACTGTTTGGCGTGCCACCCCTGAGTGTGAGCCCCAGAAATGGCTCAGTGACCTGTTTGGGA 1884
Qy 329 ----- 329
Db 1885 CCGAGGCTGACCAAGTGTGTGGCTGTGCCCCACTATAAGGACCCCTCCCTTCTGCGTGGCC 1944
Qy 329 ----- 329
Db 1945 CGCTGCCCCAGCGGTGTGAAACCTGACCTCTCTACATGCCCATCTGGAAGTTTCCAGAT 2004
Qy 329 ----- 329
Db 2005 GAGGAGGGGCGCATGCCAGCGCTTGGCCCCATCAACTGCACCCCACTCTGTGTGGACCTGGAT 2064
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:04:16 ; Search time 546.584 Seconds
(without alignments)
4313.608 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAPLLARASLSGLF.....ENPEYGLDVPAAHHHHH 555

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPO.spool/US09821883/runat.09092004.105126.6623/app_query.fasta_1.3100
-DB=Geneseq_29Jan04 -QMT=fastacp -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09821883_QCGN.1.1646 @runat.09092004.105126.6623 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003s:*
7: Geneseq2003bs:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3020	100.0	1665	4	AAD21564 Human HER
2	3005.5	99.5	1692	4	AAD21566 Human HER
3	2972	98.4	2070	4	AAD21565 Human HER
4	2957.5	97.9	2091	4	AAD21567 Human HER
5	2416	80.0	4530	2	AAT01585 Her-2/neu
6	2416	80.0	4530	2	AAT71253 Human HER
7	2416	80.0	4530	3	AAD260815 Nucleotid
8	2416	80.0	4530	4	AAD19731 Human tyr

9	2416	80.0	4530	6	ABN85585	Abn85585 Human HER
10	2416	80.0	4530	6	ABZ35012	Abz35012 Human gen
11	2416	80.0	4530	6	ABV94128	Abv94128 Breast ca
12	2416	80.0	4530	6	ABK83918	Abk83918 Human cDN
13	2416	80.0	4530	7	ACC50139	Acc50139 Breast ca
14	2416	80.0	4530	7	ABQ83856	Abq83856 Human Her
15	2416	80.0	4530	8	AAD58073	Aad58073 Human c-e
16	2416	80.0	4530	9	ADC09594	Adc09594 Her-2/neu
17	2413.5	79.9	9274	6	AAD43934	Adk43934 HER-2 tra
18	2413.5	79.9	9274	6	ABK14057	Abk14057 Human HER
19	2411	79.8	4472	3	AAA14812	Aaa14812 cDNA enco
20	2411	79.8	4473	2	ABQ76220	Abq76220 Human tum
21	2411	79.8	4473	2	AAZ31071	Aaz31071 HER-2 nuc
22	2411	79.8	4473	6	ABZ34969	Abz34969 Human gen
23	2411	79.8	4473	6	AAD38904	Aad38904 Human Her
24	2411	79.8	4473	7	ACC69999	Acc69999 Human c-e
25	2411	79.8	4473	9	ADC35148	Adc35148 Human bre
26	2411	79.8	4473	9	ADD25483	Add25483 Binding d
27	2410.5	79.8	3765	9	ADB67620	Adb67620 Human epi
28	2410.5	79.8	3768	3	AAA09455	Aaa09455 Human her
29	2410.5	79.8	3768	6	ABV78168	Abv78168 Human ERB
30	2410.5	79.8	3768	6	ABZ35744	Abz35744 Human ERB
31	2410.5	79.8	3768	6	AAD43935	Aad43935 Human HER
32	2410.5	79.8	3768	6	ABX09987	Abx09987 Human ERB
33	2410.5	79.8	3768	6	AAD43986	Aad43986 Human Her
34	2410.5	79.8	3768	6	ABK14058	Abk14058 Human HER
35	2410.5	79.8	3768	6	ABL91709	Ab191709 Human pol
36	2410.5	79.8	3768	7	ACC57649	Acc57649 Human pro
37	2406.5	79.7	9274	4	AAF24297	Aaf24297 HER2 tran
38	2405.5	79.7	3768	2	AAI40739	Aai40739 HER-2/neu
39	2405.5	79.7	3768	2	AAH01912	Aah01912 Human HER
40	2405.5	79.7	3768	5	AAH23392	Aah23392 Human HER
41	2405.5	79.7	3768	6	AAD32743	Aad32743 Human Her
42	2405.5	79.7	3768	6	ABK10730	Abk10730 Human Her
43	2405.5	79.7	3768	6	ABA92250	Abag92250 Human Her
44	2389.5	79.1	4299	2	AAQ46083	Aaq46083 Sequence
45	2386	79.0	4606	9	ADB47370	Adb47370 Human cDN

ALIGNMENTS

RESULT 1
AAD21564
ID AAD21564 standard; DNA; 1665 BP.

XX AC AAD21564;

XX DT 28-JAN-2002 (first entry)

XX DE Human HER500 fusion DNA construct.

XX KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human;
KW HER-2 protein; HER500 fusion DNA; ds.

XX OS Homo sapiens.

XX OS Synthetic.

XX OS Chimeric.

XX FH Key Location/Qualifiers
XX CDS 1..1665

XX FT /product= "Human HER500 fusion protein construct
XX FT comprising human PAP signal sequence, mature PAP protein,
XX FT an Ala Arg linker, human HER-2 signal sequence, mature
XX FT HER-2 membrane distal extracellular and intracellular
XX FT domains and a C-terminal tag"
XX FT /note= "CDS does not include stop codon"
XX FT /partial

```
PN WO200174855-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010515.
XX PF
XX 30-MAR-2000; 2000US-0193504P.
XX PR
XX (DEND-) DENDREON CORP.
XX PA
XX Laus R, Vidovic D, Graddis T;
XX PI
XX WPI; 2001-662965/76.
XX DR
XX P-PSDB; AAE13108.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain
XX PT of HER-2 and an antigen elicits an immune response to the antigen and is
XX PT useful for the treatment of associated cancer associated.
XX
XX PS Example 3; Page 27; 59pp; English.
XX
XX The invention relates to immunostimulatory fusion proteins (IPF) and
XX CC nucleic acid molecules encoding such proteins. The IPFs comprise a
XX CC polypeptide antigen component and an immunostimulatory component derived
XX CC from the intracellular domain of HER-2 protein which is effective to
XX CC elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
XX CC immune response to the antigen. IPF or superactivated dendritic cells are
XX CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
XX CC associated with a particularly antigen. The present sequence is HER500
XX CC fusion DNA construct which comprises DNA molecules encoding human PAP
XX CC signal sequence, mature PAP protein, an Ala Arg linker, human HER-2
XX CC signal sequence, mature HER-2 membrane distal extracellular and
XX CC intracellular domains and a C-terminal tag
XX
XX SQ Sequence 1665 BP; 330 A; 568 C; 465 G; 302 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,87e-140 Length: 1665
Score: 3020.00 Matches: 555
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-821-883-1 (1-555) x RAD21564 (1-1665)

Qy 1 MetArgAlaProLeuLeuLeuAlaAArgAlaAlaSerLeuSerLeuGlyPheLeuPhe 20
Db 1 ATGAGAGCTGCACCCCTCCCTCTGCGCAGGCGCAGCAAGCCCTTAGCCCTTGGCTTTGTTT 60

Qy 21 LeuLeuPhePheTTPLeuAspArgSerValLeuAlaLysGluLeuAlaAArgGlyAlaAla 40
Db 61 CTGCTTTTCTGGCTAGACCGAAGTGTAAGTACCCCAAGAGATGGCGCGCGGCGCG 120

Qy 41 SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr 60
Db 121 TCGACCCAGTGTGCACCGCAGACATGATGCTGGCTCCCTGCCAGTCCCGAGACC 180

Qy 61 HisLeuAspMetLeuArgHisLeuTyrGlnGlyCysGlnValValGlnGlyAsnLeuGlu 80
Db 181 CACCTGACATGCTCCGCCACCTCTACAGGGCTGCCAGGTGGTGGTGCAGGGAAACCTGGAA 240

Qy 81 LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspPheLeuValGln 100
Db 241 CTACCTACCTGCGCCACCAATGCCAGCTGTGCTTCTCTGAGGATATCCAGGAGGTGCAG 300

Qy 101 GlyTyrValLeuLeuLeuAlaHisAsnGlnValArgGlnValProLeuGlnArgLeuArgile 120
Db 301 GGCTACGTGCTCATCGCTCACACCAAGTGGGAGGTGCCACTGCCAGAGGCTGGGGATT 360

Qy 121 ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp 140
Db 361 GTGCGAGGCCACCCAGCTCTTTTGAGGACAACTATGCTCCCTGGCGCTGTAGCAATGGAGAC 420
```

```
Qy 141 ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyGlyLeuArgGluLeuGln 160
Db 421 CCCTGTGACATATACACCCCTCTGTCAAGGGGCTCCCGAGAGGCTCGGGAGCTGCAG 480

Qy 161 LeuArgSerLeuThrGluLeuLeuLysGlyValLeuLeuGlnArgAsnProGlnLeu 180
Db 481 CTTTGAAGCCTCACAGAGATCTTTGAAGAGGGGTCTTGATCCAGCGGAAACCCAGCTC 540

Qy 181 CysTyrGlnAspThrIleLeuTyrLysAspIlePheHisLysAsnGlnLeuAlaLeu 200
Db 541 TGCTACAGGACACGATTTTGTGAAGAGACATCTTCACAAAGAACCAACAGCTGGCTCTC 600

Qy 201 ThrLeuLeuAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly 220
Db 601 ACCTGATAGACACCAACCCCTCTCGGGCTGCCACCCCTGTTCTCGATGTGTAAAGGC 660

Qy 221 SerArgCysTyrGlyGluSerGluAspCysGlnSerLeuThrArgThrValCysAla 240
Db 661 TCCCGCTGTGGGAGAGAGATTCTGAGGATTGTGAGAGCCTGACGCGCACTGTCTGTGCC 720

Qy 241 GlyGlyCysAlaArgCysLysGlyProLeuProThrAspCysCysHisGluGlnCysAla 260
Db 721 GGTGGCTGTGCCCTGTCAAGGGGCGCCTGCCACTGACTGCTGCCATGAGCAGTGTGCT 780

Qy 261 AlaGlyCysThrGlyProLysHisSerAspCysLeuAlaCysLeuHisPheAsnHisSer 280
Db 781 GCCGGCTGCAGGGGCGCCAGCACTCTGACTGCTGGCTGCCCTGCCATTCACCCAGT 840

Qy 281 GlyIleCysGluLeuHisCysProAlaLeuValThrTyrAsnThrAspThrPheGluSer 300
Db 841 GGCATCTGTGAGCTGCACCTGCCAGGCCCTGGTCACTACCAACACAGACAGCTTTAGTCC 900

Qy 301 MetProAsnProGluGlyArgTyrThrPheGlyAlaSerCysValThrAlaCysProTyr 320
Db 901 ATGCCCAATCCGAGGGCGGTATACATTCGGCGCCAGCTGTGTGACTGCTGTGCCCTAC 960

Qy 321 AsnTyrLeuSerThrAspValGlySerGlyValAlaGlyGlyMetValHisHisArgHisArg 340
Db 961 AACTACTCTTCTACGACGCTGGGATCGGGGCGCTGGGGGCGATGTTCCACACAGCAGCCGC 1020

Qy 341 SerSerSerThrArgSerGlyGlyAspLeuThrLeuGlyLeuGluProSerGluGlu 360
Db 1021 AGCTCATCTACAGAGAGTGGCGTGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAG 1080

Qy 361 GluAlaProArgSerProLeuAlaProSerGluGlyAlaGlySerAspValPheAspGly 380
Db 1081 GAGGCCCCCAGGCTCTCCACTGGCACCCTCCGAAGGGGCTGGCTCCGATGATTTGATGGT 1140

Qy 381 AspLeuGlyMetGlyAlaAlaLysGlyLeuGlnSerLeuProThrHisAspProSerPro 400
Db 1141 GACTTGGGAATGGGGGCGCAGCCAGGGGCTGCAAAAGCTCCCCACACATGACCCAGCCCT 1200

Qy 401 LeuGlnArgTyrSerGluAspProThrValProLeuProSerGluThrAspGlyTyrVal 420
Db 1201 CTACAGCGGTACAGTGGAGACCCACAGTACCCCTGCCCTCTGAGACTGATGCTACGTT 1260

Qy 421 AlaProLeuThrCysSerProGlnProGluTyrValAsnGlnProAspValArgProGln 440
Db 1261 GCCCCCTTGACCTGACGCCCCCAGCCCTGATATGTGAACCCAGCAGATGTTTCGGCCCCAG 1320

Qy 441 ProProSerProArgGluGlyProLeuProAlaAlaArgProAlaGlyAlaThrLeuGlu 460
Db 1321 CCCCTCTGCCCGCAGAGGGGCTCTGCTGCTGCCAGCTGCTGGTGCACCTCTGGAA 1380

Qy 461 ArgAlaLysThrLeuSerProGlyLysAsnGlyValValLysAspValPheAlaPheGly 480
Db 1381 AGGGCCAAAGACTCTCTCCCGAGGGAAGATGGGGTGGTCAAAGACGTTTTTGGCTTTGGG 1440

Qy 481 GlyAlaValGluAsnProGluTyrLeuThrProGlnGlyAlaAlaProGlnProHis 500
Db 1441 GGTGCGCTGGAGAAACCCCGAGTACTTGACACCCCGAGGAGAGCTGCCCTCTCAGCCCCAC 1500
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 10, 2004, 05:41:11 ; Search time 5945.99 Seconds
(without alignments)
4045.651 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLAARASLSGLF.....ENPEYLGLDVPAAAHHHHH 555

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO/US09821883/runat_09092004_105127_6631/app_query.fasta_1.3100
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0 -LOPEXT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09821883 @CGN 1 1 16795 @runat_09092004_105127_6631 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database :
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mus.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_ats.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3020	100.0	1665	6	AX268285 Sequence
2	3005.5	99.5	1692	6	AX268287 Sequence
3	2972	98.4	2070	6	AX268286 Sequence
4	2957.5	97.9	2091	6	AX268288 Sequence
5	2416	80.0	4530	6	I21124 Sequence 9
6	2416	80.0	4530	6	I59745 Sequence 9
7	2416	80.0	4530	6	AR202597 Sequence
8	2416	80.0	4530	6	AR283481 Sequence
9	2416	80.0	4530	6	AR344811 Sequence
10	2416	80.0	4530	6	AX282577 Sequence
11	2416	80.0	4530	6	AX587649 Sequence
12	2416	80.0	4530	6	AX644071 Sequence
13	2416	80.0	4530	6	AX771418 Sequence
14	2416	80.0	4530	6	BD005474 Cellular
15	2416	80.0	4530	9	HUMHER2A
16	2413.5	79.9	9274	6	AR409602 Sequence
17	2413.5	79.9	9274	6	AX060703 Sequence
18	2411	79.8	4473	6	AR080259 Sequence
19	2411	79.8	4473	6	AR167390 Sequence
20	2411	79.8	4473	6	AR392088 Sequence
21	2411	79.8	4473	9	HSERB2R
22	2410.5	79.8	3768	6	AR409603 Sequence
23	2410.5	79.8	3768	6	AX060704 Sequence
24	2410.5	79.8	3768	6	AX467229 Sequence
25	2410.5	79.8	3768	6	AX481438 Sequence
26	2410.5	79.8	3768	6	BD224136 Novel met
27	2405.5	79.7	3768	6	AR034479 Sequence
28	2405.5	79.7	3768	6	BD267514 HER-2/neu
29	2405.5	79.7	3768	6	AX201817 Sequence
30	2405.5	79.7	3768	6	AX380923 Sequence
31	2405.5	79.7	3768	6	AX384604 Sequence
32	2405.5	79.7	3768	6	AX465456 Sequence
33	2334.5	77.3	3678	6	AX505114 Sequence
34	2129	70.5	2763	6	AX380942 Sequence
35	2129	70.5	2781	6	AX380944 Sequence
36	2124.5	70.3	3780	4	AB008451 Canis fam
37	1977	65.5	4506	10	AK129487 Mus muscu
38	1963.5	65.0	4694	10	BC046811 Mus muscu
39	1963.5	65.0	4695	10	BC053078 Mus muscu
40	1961	64.9	3771	6	BD267516 HER-2/neu
41	1961	64.9	3771	6	AX189662 Sequence
42	1961	64.9	3771	6	AX380925 Sequence
43	1958	64.8	4062	10	HANNEU
44	1957	64.8	4727	10	AY116182
45	1956.5	64.8	3955	6	BD267515 HER-2/neu

ALIGNMENTS

RESULT 1

AX268285	LOCUS	AX268285	1665 bp	DNA	linear	FAT 29-OCT-2001
DEFINITION	Sequence 6 from Patent WO0174855.					
ACCESSION	AX268285					
VERSION	AX268285.1	GI:16541537				
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	artificial sequences.					
REFERENCE	1					
AUTHORS	Laus, R., Vidovic, D. and Graddis, T.					
TITLE	Compositions and methods for dendritic cell-based immunotherapy					
JOURNAL	Patent: WO 0174855-A 6 11-OCT-2001;					
FEATURES	DENDREON CORPORATION (US)					
source	Location/Qualifiers					
	1..1665					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="HER500 construct"					
ORIGIN						
Alignment Scores:						
Pred. No.:	2,44e-110	Length:	1665			
Score:	3020.00	Matches:	555			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			
US-09-821-883--1 (1-555) x AX268285 (1-1665)						
Qy	1	MetArgAlaAlaProLeuLeuLeuAlaArgAlaAlaSerLeuSerLeuGlyPheLeuPhe	20			
Db	1	ATGAGAGCTGCACCCCTCTCTCGGCGAGGAGCAGCAAGCCTTAGCCTTTGTTT	60			
Qy	21	LeuLeuPhePheTrpLeuAspArgSerValLeuAlaLysGluLeuAlaAargGlyAlaAla	40			
Db	61	CTGCTTTTCTGGCTAGACCGAAGTGTAAGCAAGGAGTGGCGCGGGCGCG	120			
Qy	41	SerThrGlnValCysThrGlyThrAspMetLysLeuArgLeuProAlaSerProGluThr	60			
Db	121	TCGACCCCAAGTGTGCACCGGACAGACATGAAGCTGCGGCTCCCTGCCAGTCCCGAGACC	180			
Qy	61	HisLeuAspMetLeuArgHisLeuTrpGlnGlyCysGlnValValGlnGlyAsnLeuGlu	80			
Db	181	CACCTGACATGCTCGCCACCTCTACCAAGGCTGCGAGGTGGCGAGGAACTGGNA	240			
Qy	81	LeuThrTyrLeuProThrAsnAlaSerLeuSerPheLeuGlnAspIleGlnValGln	100			
Db	241	CTCACCTACCTGCCACCAATGCGACCTGTCTTCTGTCAGATATCCAGGAGGTGCAG	300			
Qy	101	GlyTyrValLeuIleAlaHisGlnValArgGlnValProLeuGlnArgLeuArgIle	120			
Db	301	GGCTAGCTGCTATCGTCCACCAAGTCAGGACAGTCCCACTGCAGAGGCTGCGGATT	360			
Qy	121	ValArgGlyThrGlnLeuPheGluAspAsnTyrAlaLeuAlaValLeuAspAsnGlyAsp	140			
Db	361	GTGCGAGGACCCAGCTCTTTGAGGACAACTATGCCCCCTGGCCGTGTAGACAATGGAGAC	420			
Qy	141	ProLeuAsnAsnThrThrProValThrGlyAlaSerProGlyLeuArgGluLeuGln	160			
Db	421	CCGCTGNACATACCAACCCCTGTACAGGGGCTCCCGAGGAGGCTGCGGAGGTGCAG	480			
Qy	161	LeuArgSerLeuThrGluIleLeuLysGlyValLeuIleGlnArgAsnProGlnLeu	180			
Db	481	CTTCGAAGCCTCACAGAGATCTTGAAGGAGGGGTCTTGATCCAGCGGAAACCCCGAGTC	540			
Qy	181	CysTyrGlnAspThrIleLeuTrpLysAspIlePheHisLysAsnAsnGlnLeuAlaLeu	200			
Db	541	TGCTACAGAGACAGATTTTGTGGAGGAGCAATCTTCCACAAAGAACACACGCTGGCTCTC	600			
Qy	201	ThrLeuIleAspThrAsnArgSerArgAlaCysHisProCysSerProMetCysLysGly	220			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:28:07 ; Search time 55.1456 Seconds
(without alignments)
3175.456 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLAARASLSGLF.....ENPEYGLDVPAAHFHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2124.5	70.3	1259	6	O18735 canis famil
2	1951	64.6	1259	11	Q8K3F9
3	1609.5	53.3	419	4	Q9UK79
4	1376.5	45.6	711	11	Q80V89
5	1005.5	33.3	881	11	Q8C0E7
6	969	32.1	367	11	Q8R2X1
7	906	30.0	165	4	Q14256
8	900.5	29.8	412	4	Q8WYV0
9	861.5	28.5	431	13	Q7SY19
10	754.5	25.0	527	13	Q90836
11	723	23.9	149	6	Q9BG66
12	717.5	23.8	643	11	Q9ERV6
13	717.5	23.8	655	11	Q9WVFS
14	717.5	23.8	1210	11	Q9EP98
15	714.5	23.7	1209	11	Q9QX70
16	708.5	23.5	1209	6	Q8M1L8

ALIGNMENTS

RESULT 1

ID	O18735	PRELIMINARY;	PRT;	1259 AA.
AC	O18735;			
DT	01-JAN-1998 (TrEMBLrel. 05, Created)			
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Erbb-2			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Yokota H.;			
RT	"CDNA cloning of erbb-2 from canine mammary gland.";			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBSJ databases.			
DR	EMBL; AB008451; BAA23127.1; -			
DR	HSSP; P11362; 1FGK			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0005534; F:ATP binding; IEA.			
DR	GO; GO:0005509; F:calcium ion binding; IEA.			
DR	GO; GO:0005006; F:epidermal growth factor; IEA.			
DR	GO; GO:0016740; F:transferase activity; IEA.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
DR	GO; GO:0007189; P:transmembrane receptor protein tyrosine kin. . . ; IEA.			
DR	InterPro; IPR000494; EGFR_L_domain.			
DR	InterPro; IPR006211; Furin-like.			
DR	InterPro; IPR006212; Furin repeat.			
DR	InterPro; IPR009030; Grow_Fac_recep.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR001245; Tyr_kinase.			
DR	InterPro; IPR008266; Tyr_kinase_AS.			
DR	InterPro; IPR004019; YLP motif.			
DR	Pfam; PF00757; Furin-like; 1.			
DR	Pfam; PF00065; pkinase; 1.			
DR	Pfam; PF01030; Recep_L_domain; 2.			
DR	Pfam; PF02757; YLP; 2.			

Q9EE60 rattus norv
Q78ZF7 brachydanio
Q9BUD7 homo sapien
Q9YH40 xiphophorus
Q8AW81 brachydanio
Q9PEH2 gallus gall
P79754 fugu rubrip
Q8M1W0 drosophila
Q9B1H9 anopheles g
Q8ENZ2 drosophila
Q9W6F6 gallus gall
Q8BG64 oryctolagus
Q9B965 oryctolagus
Q23821 caenorhabdi
Q26569 schistosoma
Q26566 schistosoma
Q26567 schistosoma
Q26568 schistosoma
Q9Y1X8 ephydatia f
Q8EMD7 echinococcu
Q9U5A8 bombyx mori
Q9NJV5 biophalar
Q8T0W6 echinococcu
Q88459 mus musculu
Q8UW83 paralichthy
Q93457 scophthalmu
Q9VD94 drosophila
Q73798 xenopus lae
Q8UW84 paralichthy

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:54 ; Search time 10.1875 Seconds
(without alignments)
2836.691 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLAARASLSGLF.....ENPEYGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2405.5	79.7	1255	1 ERB2 HUMAN	P04626 homo sapien
2	1947.5	64.5	1257	1 ERB2 RAT	P06494 rattus norv
3	1935	64.1	1254	1 ERB2 MESAU	Q60553 mesocricetu
4	755.5	25.0	703	1 EGFR CHICK	P13387 gallus gall
5	721	23.9	1210	1 EGFR HUMAN	P00533 homo sapien
6	721	23.9	1308	1 ERB4 HUMAN	Q15303 homo sapien
7	717.5	23.8	1210	1 EGFR MOUSE	Q01279 mus musculu
8	717	23.7	1308	1 ERB4 RAT	Q62956 rattus norv
9	697	23.1	1342	1 ERB3 HUMAN	P21860 homo sapien
10	668	22.1	1339	1 ERB3 RAT	Q62799 rattus norv
11	655.5	21.7	1167	1 XMRK XIPMA	P13388 xiphophorus
12	536	17.7	1426	1 EGFR DROME	P04412 drosophila
13	404.5	13.4	1367	1 L723 CAEEL	P24348 caenorhabdi
14	317.5	10.5	1363	1 ILPR BRAUA	Q02466 brachiosto
15	293	9.7	245	1 ERB2 MOUSE	P70424 mus musculu
16	273.5	9.1	1477	1 HTK7 HYDAT	Q25197 hydra atten
17	270.5	9.0	2146	1 INSR DROME	P09208 drosophila
18	256.5	8.5	1382	1 INSR HUMAN	P06213 homo sapien
19	232.5	8.4	1372	1 INSR MOUSE	P5208 mus musculu
20	252	8.3	1607	1 MIPR LYMET	Q35410 lymanaea sta
21	251.5	8.3	1300	1 IRR MOUSE	Q04710 mus musculu
22	249	8.2	1383	1 INSR RAT	P15127 rattus norv
23	248.5	8.2	1297	1 IRR HUMAN	P14616 homo sapien
24	246.5	8.2	1300	1 IRR CAVPO	P14617 cavia porce
25	246	8.1	1367	1 IGIR HUMAN	P08069 homo sapien
26	235.5	7.8	1370	1 IGIR RAT	P24062 rattus norv
27	233.5	7.7	1373	1 IGIR MOUSE	Q60751 mus musculu
28	228.5	7.6	581	1 IRR RAT	Q64716 rattus norv
29	207	6.9	1390	1 INSR AEDA	Q93105 aedes aegyp
30	168	5.6	386	1 PPAP HUMAN	P53309 homo sapien
31	155	5.1	1321	1 IRS2 MOUSE	P81122 mus musculu
32	151.5	5.0	707	1 SPQO HUMAN	P23246 homo sapien
33	151.5	5.0	5262	1 MLL2 HUMAN	O14686 homo sapien

ALIGNMENTS

RESULT 1

ID	ERB2_HUMAN	STANDARD;	PRT;	1255 AA.
AC	P04626;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	13-AUG-1987 (Rel. 05, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Receptor protein-tyrosine kinase erbB-2 precursor (EC 2.7.1.112)			
DE	(p185erbB2) (NEU proto-oncogene) (C-erbB-2) (Tyrosine kinase-type cell			
DE	surface receptor HER2) (MLN 19).			
GN	ERBB2 OR HER2 OR NGL OR NEU.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96118663; PubMed=3003577;			
RA	Yamamoto T., Ikawa S., Akiyama T., Semba K., Nomura N., Miyajima N.,			
RA	Saito T., Toyoshima K.;			
RT	"Similarity of protein encoded by the human c-erb-B-2 gene to			
RT	epidermal growth factor receptor.";			
RL	Nature 319:230-234(1986).			
[2]				
RP	SEQUENCE FROM N.A., AND VARIANT ALA-1170.			
RX	MEDLINE=86070181; PubMed=2999974;			
RA	Coussens L., Yang-Feng T.L., Liao Y.C., Chen E., Gray A.,			
RA	McGrath J., Seeburg P.H., Libermann T.A., Schlessinger J.,			
RA	Francie U., Levinson A., Ullrich A.;			
RT	"Tyrosine kinase receptor with extensive homology to EGF receptor			
RT	shares chromosomal location with neu oncogene.";			
RL	Science 230:1132-1139(1985).			
[3]				
RP	SEQUENCE FROM N.A., AND VARIANTS CYS-452; VAL-655 AND ALA-1170.			
RX	Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,			
RA	Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,			
RA	Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;			
RL	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.			
[4]				
RP	SEQUENCE OF 737-1031 FROM N.A.			
RX	MEDLINE=86016729; PubMed=295967;			
RA	Semba K., Kamata N., Toyoshima K., Yamamoto T.;			
RT	"A v-erbB-related protooncogene, c-erbB-2, is distinct from the			
RT	c-erbB-1/epidermal growth factor-receptor gene and is amplified in a			
RT	human salivary gland adenocarcinoma.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:6497-6501(1985).			
[5]				
RP	VARIANTS VAL-654 AND VAL-655.			
RX	MEDLINE=93194196; PubMed=8095488;			
RA	Ehsani A., Low J., Wallace R.B., Wu A.M.;			
RT	"Characterization of a new allele of the human ERBB2 gene by allele-			
RT	specific competition hybridization.";			
RL	Genomics 15:426-429(1993).			
CC	-1- FUNCTION: Essential component of a neurotrophin-receptor complex,			
CC	although neurotrophins do not interact with it alone. GP30 is a			
CC	potential ligand for this receptor. Not activated by EGF, TGF-			

34	141.5	4.7	830	1	SREC_HUMAN	Q14162 homo sapien
35	141.5	4.7	2442	1	CBP_HUMAN	Q92793 homo sapien
36	140	4.6	863	1	ML13_HUMAN	Q9N3F8 homo sapien
37	140	4.6	1696	1	PCK5_BRACL	Q9N155 branchiosto
38	139	4.6	634	1	ERBB2_ALV	P00534 avian leuko
39	139	4.6	870	1	SRC2_HUMAN	Q96966 homo sapien
40	138.5	4.6	553	1	ODO2_MYCTU	Q10381 mycobacteri
41	137.5	4.6	620	1	EXTN2_TOBAC	P13983 nicotiana t
42	136	4.5	331	1	PRP1_HUMAN	P04280 homo sapien
43	133.5	4.4	886	1	SM6B_MOUSE	O54951 mus musculu
44	131.5	4.4	1185	1	DRPL_HUMAN	P54259 homo sapien
45	130.5	4.3	833	1	SRC2_MOUSE	P59222 mus musculu

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:29:10 ; Search time 17.7175 Seconds
(without alignments)
3013.200 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLIARRASLSGLFLP.....ENPEYLGLDVPAAAHHHHHH 555
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405.5	79.7	1255	1 A24571	protein-tyrosine k
2	1950.5	64.6	1260	1 TVRTNU	protein-tyrosine k
3	1935	64.1	1254	2 I48161	p-185 precursor -
4	755.5	25.0	1223	1 TVCHLV	epidermal growth f
5	754.5	25.0	527	2 A42032	epidermal growth f
6	721	23.9	1210	1 GQHUE	epidermal growth f
7	721	23.9	1308	2 A47253	epidermal growth f
8	717.5	23.8	1210	2 A53183	epidermal growth f
9	714.5	23.7	644	2 A36325	epidermal growth f
10	697	23.1	1342	2 A36223	kinase-related tra
11	662	21.9	1339	2 JC4387	epidermal growth f
12	655.5	21.7	1166	1 S06142	protein-tyrosine k
13	536	17.7	843	2 A27131	epidermal growth f
14	404.5	13.4	1323	2 E88257	protein let-23 (im
15	404.5	13.4	1374	2 S70712	protein-tyrosine k
16	390.5	12.9	1369	2 S70713	protein-tyrosine k
17	375.5	12.4	1330	1 GQFFE	epidermal growth f
18	341	11.3	366	2 D45558	epidermal growth f
19	341	11.3	1717	1 A45558	epidermal growth f
20	331	11.0	333	2 B45558	epidermal growth f
21	331	11.0	342	2 C45558	epidermal growth f
22	317.5	10.5	1363	2 T43220	insulin-like growth
23	273.5	9.1	1477	2 T18534	protein-tyrosine k
24	270.5	9.0	2101	2 S57245	insulin receptor (
25	270.5	9.0	2148	1 A56081	insulin receptor -
26	256.5	8.5	1382	1 INHUR	insulin receptor p
27	252.5	8.4	1372	2 A34157	insulin receptor p
28	252	8.3	1607	2 T43212	insulin-like growth
29	249	8.2	1383	2 A36080	insulin receptor p

30 246.5 8.2 1300 2 A36502 insulin receptor-r
31 246 8.1 1367 1 IGHURI insulin-like growth
32 238 7.9 540 2 B47417 insulin receptor-r
33 235.5 7.8 183 2 JH0803 tyrosine kinase re
34 235.5 7.8 1371 2 A33837 insulin-like growth
35 231.5 7.7 1268 2 B36502 insulin receptor-r
36 223.5 7.4 329 2 A48805 insulin-like growth
37 207 6.9 1390 2 T30346 insulin receptor -
38 188 6.2 1846 2 T42047 insulin receptor h
39 168 5.6 386 1 JH0610 acid phosphatase (

ALIGNMENTS

RESULT 1

A24571
protein-tyrosine kinase (EC 2.7.1.112) erbB2 precursor - human
N/Alternate names: c-erb-B-2 protein precursor; kinase-related transforming protein erbB
C/Species: Homo sapiens (man)
C/Date: 25-Oct-1987 #sequence revision 06-Dec-1996 #text_change 11-Jun-1999
C/Accession: A24571; A25491; A44188; B44188; I59509; I57622
R/Yamamoto, T.; Ikawa, S.; Akiyama, T.; Semba, K.; Nomura, N.; Miyajima, N.; Saito, T.;
Nature 319, 230-234, 1986
A/Title: Similarity of protein encoded by the human c-erb-B-2 gene to epidermal growth f
A/Reference number: A24571; MUID:86118663; PMID:3003577
A/Accession: A24571
A/Molecule type: mRNA
A/Residues: 1-1255 <YAM>
A/Cross-references: GB:X03363; NID:g31197; PIDN:CAA27060.1; PID:g31198
R/Semba, K.; Kamata, N.; Toyoshima, K.; Yamamoto, T.
Proc. Natl. Acad. Sci. U.S.A. 82, 6497-6501, 1985
A/Title: A v-erbB-related protooncogene, c-erbB-2, is distinct from the c-erbB-1/epiderm
A/Reference number: A25491; MUID:86016729; PMID:2995967
A/Accession: A25491
A/Molecule type: DNA
A/Residues: 737-1031 <SEM>
A/Cross-references: GB:M11767; NID:g182163; PIDN:AAA35808.1; PID:g553282
R/Coussens, L.; Yang-Feng, T.L.; Liao, Y.C.; Chen, E.; Gray, A.; McGrath, J.; Seeburg, P.
Science 230, 1132-1139, 1985
A/Title: Tyrosine kinase receptor with extensive homology to EGF receptor shares chromos
A/Reference number: A44188; MUID:86070181; PMID:2999974
A/Accession: A44188
A/Molecule type: DNA
A/Residues: 740-910 <COU1>
A/Cross-references: GB:M12036; NID:g183988; PIDN:AAA35978.1; PID:g183989
A/Accession: B44188
A/Molecule type: mRNA
A/Residues: 1-517; RALL, 522, 'S', 524-654, 'V', 656-1169, 'A', 1171-1255 <COU2>
A/Cross-references: GB:M11730; NID:g183986
R/King, C.R.; Kraus, M.H.; Aaronson, S.A.
Science 229, 974-976, 1985
A/Title: Amplification of a novel v-erbB-related gene in a human mammary carcinoma.
A/Reference number: I59509; MUID:85272597; PMID:2992089
A/Accession: I59509
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 832-909 <REX>
A/Cross-references: GB:I29395; NID:g459807; PIDN:AAA35809.1; PID:g459808
R/Tal, M.; King, C.R.; Kraus, M.H.; Ullrich, A.; Schlessinger, J.; Givol, D.
Mol. Cell. Biol. 7, 2597-2601, 1987
A/Title: Human HER2 (neu) promoter: evidence for multiple mechanisms for transcriptional
A/Reference number: I57622; MUID:87286898; PMID:3039351
A/Accession: I57622
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-191 <TAL>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:41:37 ; Search time 58.6891 Seconds
(without alignments)
3032.635 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSIGFLF.....ENPEYGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3020	100.0	555	US-09-821-883-1	Sequence 1, Appli
2	3005.5	99.5	564	US-09-821-883-3	Sequence 3, Appli
3	2972	98.4	690	US-09-821-883-2	Sequence 2, Appli
4	2957.5	97.9	697	US-09-821-883-4	Sequence 4, Appli
5	2573.5	85.2	919	US-09-854-356-6	Sequence 6, Appli
6	2410.5	79.8	1255	US-09-811-123-9	Sequence 9, Appli
7	2410.5	79.8	1255	US-09-811-115-3	Sequence 3, Appli
8	2410.5	79.8	1255	US-09-769-508-2	Sequence 2, Appli
9	2410.5	79.8	1255	US-09-984-092-4	Sequence 4, Appli
10	2410.5	79.8	1255	US-10-177-293-126	Sequence 126, App
11	2410.5	79.8	1255	US-10-207-498-6	Sequence 6, Appli
12	2410.5	79.8	1255	US-10-338-730-2	Sequence 2, Appli
13	2410.5	79.8	1255	US-10-322-892-4	Sequence 4, Appli
14	2410.5	79.8	1255	US-10-272-437A-28	Sequence 28, Appl
15	2410.5	79.8	1255	US-10-117-937-594	Sequence 594, App

ALIGNMENTS

RESULT 1

US-09-821-883-1
; Sequence 1, Application US/09821883
; Patent No. US20020061310A1
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500 construct
US-09-821-883-1

Query Match 100.0%; Score 3020; DB 9; Length 555;

Best Local Similarity 100.0%; Pred. No. 4.5e-203;

Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLARASLSIGFLFLFFWLDKRSVLAKELARGAASQVCTGTGDKMLRASPPT 60

DB 1 MRAAPLLARASLSIGFLFLFFWLDKRSVLAKELARGAASQVCTGTGDKMLRASPPT 60

QY 61 HUDMLRHLYQGCQVQGNLELYLPTNASLSFLQDIEQVGVVLAHNVQVPLQRLRI 120

DB 61 HUDMLRHLYQGCQVQGNLELYLPTNASLSFLQDIEQVGVVLAHNVQVPLQRLRI 120

QY 121 VRGTQLFEDNYALVDNGDPLNNTTPTVTGASFGGLRELQRLSLTEILKGGVL1QRNPOL 180

Sequence 36, Appli

Sequence 4, Appli

Sequence 126, Appli

Sequence 1, Appli

Sequence 2, Appli

Sequence 3, Appli

Sequence 4, Appli

Sequence 5, Appli

Sequence 6, Appli

Sequence 7, Appli

Sequence 8, Appli

Sequence 9, Appli

Sequence 10, Appli

Sequence 11, Appli

Sequence 12, Appli

Sequence 13, Appli

Sequence 14, Appli

Sequence 15, Appli

Sequence 16, Appli

Sequence 17, Appli

Sequence 18, Appli

Sequence 19, Appli

Sequence 20, Appli

Sequence 21, Appli

Sequence 22, Appli

Sequence 23, Appli

Sequence 24, Appli

Sequence 25, Appli

Sequence 26, Appli

Sequence 27, Appli

Sequence 28, Appli

Sequence 29, Appli

Sequence 30, Appli

Sequence 31, Appli

Sequence 32, Appli

Sequence 33, Appli

Sequence 34, Appli

Sequence 35, Appli

Sequence 36, Appli

Sequence 37, Appli

Sequence 38, Appli

Sequence 39, Appli

Sequence 40, Appli

Sequence 41, Appli

Sequence 42, Appli

Sequence 43, Appli

Sequence 44, Appli

Sequence 45, Appli

Sequence 46, Appli

Sequence 47, Appli

Sequence 48, Appli

Sequence 49, Appli

Sequence 50, Appli

Sequence 51, Appli

Sequence 52, Appli

Sequence 53, Appli

Sequence 54, Appli

Sequence 55, Appli

Sequence 56, Appli

Sequence 57, Appli

Sequence 58, Appli

Sequence 59, Appli

Sequence 60, Appli

Sequence 61, Appli

Sequence 62, Appli

Sequence 63, Appli

Sequence 64, Appli

Sequence 65, Appli

Sequence 66, Appli

Sequence 67, Appli

Sequence 68, Appli

Sequence 69, Appli

Sequence 70, Appli

Sequence 71, Appli

Sequence 72, Appli

Sequence 73, Appli

Sequence 74, Appli

Sequence 75, Appli

Sequence 76, Appli

Sequence 77, Appli

Sequence 78, Appli

Sequence 79, Appli

Sequence 80, Appli

Sequence 81, Appli

Sequence 82, Appli

Sequence 83, Appli

Sequence 84, Appli

Sequence 85, Appli

Sequence 86, Appli

Sequence 87, Appli

Sequence 88, Appli

Sequence 89, Appli

Sequence 90, Appli

Sequence 91, Appli

Sequence 92, Appli

Sequence 93, Appli

Sequence 94, Appli

Sequence 95, Appli

Sequence 96, Appli

Sequence 97, Appli

Sequence 98, Appli

Sequence 99, Appli

Sequence 100, Appli

Sequence 101, Appli

Sequence 102, Appli

Sequence 103, Appli

Sequence 104, Appli

Sequence 105, Appli

Sequence 106, Appli

Sequence 107, Appli

Sequence 108, Appli

Sequence 109, Appli

Sequence 110, Appli

Sequence 111, Appli

Sequence 112, Appli

Sequence 113, Appli

Sequence 114, Appli

Sequence 115, Appli

Sequence 116, Appli

Sequence 117, Appli

Sequence 118, Appli

Sequence 119, Appli

Sequence 120, Appli

Sequence 121, Appli

Sequence 122, Appli

Sequence 123, Appli

Sequence 124, Appli

Sequence 125, Appli

Sequence 126, Appli

Sequence 127, Appli

Sequence 128, Appli

Sequence 129, Appli

Sequence 130, Appli

Sequence 131, Appli

Sequence 132, Appli

Sequence 133, Appli

Sequence 134, Appli

Sequence 135, Appli

Sequence 136, Appli

Sequence 137, Appli

Sequence 138, Appli

Sequence 139, Appli

Sequence 140, Appli

Sequence 141, Appli

Sequence 142, Appli

Sequence 143, Appli

Sequence 144, Appli

Sequence 145, Appli

Sequence 146, Appli

Sequence 147, Appli

Sequence 148, Appli

Sequence 149, Appli

Sequence 150, Appli

Sequence 151, Appli

Sequence 152, Appli

Sequence 153, Appli

Sequence 154, Appli

Sequence 155, Appli

Sequence 156, Appli

Sequence 157, Appli

Sequence 158, Appli

Sequence 159, Appli

Sequence 160, Appli

Sequence 161, Appli

Sequence 162, Appli

Sequence 163, Appli

Sequence 164, Appli

Sequence 165, Appli

Sequence 166, Appli

Sequence 167, Appli

Sequence 168, Appli

Sequence 169, Appli

Sequence 170, Appli

Sequence 171, Appli

Sequence 172, Appli

Sequence 173, Appli

Sequence 174, Appli

Sequence 175, Appli

Sequence 176, Appli

Sequence 177, Appli

Sequence 178, Appli

Sequence 179, Appli

Sequence 180, Appli

Sequence 181, Appli

Sequence 182, Appli

Sequence 183, Appli

Sequence 184, Appli

Sequence 185, Appli

Sequence 186, Appli

Sequence 187, Appli

Sequence 188, Appli

Sequence 189, Appli

Sequence 190, Appli

Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240
QY 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
QY 301 MPNPEGRVTFGASCVTACPYNYLSTDVSGAGGMVHRRSSSTRSGGGDLTLGLEPSEE 360
Db 301 MPNPEGRVTFGASCVTACPYNYLSTDVSGAGGMVHRRSSSTRSGGGDLTLGLEPSEE 360
QY 361 EAPRSLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVPLPSETDGVV 420
Db 361 EAPRSLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTVPLPSETDGVV 420
QY 421 APLTCSPOPEYVNDVRRPQPSREGPLPAARPAAGATLERAKTILSPGKGVVVDVFAFG 480
Db 421 APLTCSPOPEYVNDVRRPQPSREGPLPAARPAAGATLERAKTILSPGKGVVVDVFAFG 480
QY 481 GAVENPEYLTPOGGAAPQHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
Db 481 GAVENPEYLTPOGGAAPQHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKGTPTAENPEY 540
QY 541 LGLDVPAHAHHHHH 555
Db 541 LGLDVPAHAHHHHH 555

RESULT 2

US-09-821-883-3
; Sequence 3, Application US/09821883
; Patent No. US20020061310A1

GENERAL INFORMATION:

; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 564
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500* construct

US-09-821-883-3

Query Match 99.5%; Score 3005.5; DB 9; Length 564;
Best Local Similarity 98.4%; Pred. No. 4.8e-202;
Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGASTQVCTGTDKMLRLPASPET 60
Db 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGASTQVCTGTDKMLRLPASPET 60
QY 61 HLDMLRLHYQGCQVQVQGNLELYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLRI 120
Db 61 HLDMLRLHYQGCQVQVQGNLELYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLRI 120
QY 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180

QY 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240
QY 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
Db 241 GGCARCKGPLPTDCCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVTYNTDTPES 300
QY 301 MPNPEGRVTFGASCVTACPYNYLSTDVGS-----GAGGMVHRRSSSTRSGGGDL 351
Db 301 MPNPEGRVTFGASCVTACPYNYLSTDVGSASIIINEFKLGAGGMVHRRSSSTRSGGGDL 360
QY 352 TLGLEPSEEEAAPSRLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTV 411
Db 352 TLGLEPSEEEAAPSRLAPSEAGSDVFDGLGMAAGLQSLPHTDPSPLQRYSEDPTV 420
QY 412 LPSETDGVVAPLTCSPQPEYVNDVRRPQPSREGPLPAARPAAGATLERAKTILSPGKNG 471
Db 412 LPSETDGVVAPLTCSPQPEYVNDVRRPQPSREGPLPAARPAAGATLERAKTILSPGKNG 480
QY 472 VVKOVFAFGGAVENPEYLTPOGGAAPQHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 531
Db 472 VVKOVFAFGGAVENPEYLTPOGGAAPQHPPPAFSPAFDNLVYWDQDPPPERGAPPSTFKG 540
QY 532 TPTAENPEYLTGIDVPAHAHHHHH 555
Db 532 TPTAENPEYLTGIDVPAHAHHHHH 555
QY 541 TPTAENPEYLTGIDVPAHAHHHHH 564
Db 541 TPTAENPEYLTGIDVPAHAHHHHH 564

RESULT 3

US-09-821-883-2

; Sequence 2, Application US/09821883

; Patent No. US20020061310A1

GENERAL INFORMATION:

; APPLICANT: Laus, Reiner
; APPLICANT: Vidovic, Damir
; APPLICANT: Graddis, Thomas
; TITLE OF INVENTION: Compositions and Methods for Dendritic
; TITLE OF INVENTION: Cell-Based Immunotherapy
; FILE REFERENCE: 7636-0022.30
; CURRENT APPLICATION NUMBER: US/09/821,883
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,504
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HER500-hgm-CSF construct

US-09-821-883-2

Query Match 98.4%; Score 2972; DB 9; Length 690;
Best Local Similarity 100.0%; Pred. No. 1.4e-199;
Matches 549; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGASTQVCTGTDKMLRLPASPET 60
Db 1 MRAAPLLIARAASLSLGLFLLFFWLDRLSVLAKELARGASTQVCTGTDKMLRLPASPET 60
QY 61 HLDMLRLHYQGCQVQVQGNLELYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLRI 120
Db 61 HLDMLRLHYQGCQVQVQGNLELYLPTNASLSFLQDIQEVQGVVLIHNVQVPLQRLRI 120
QY 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
Db 121 VGGTQLFEDNYALVLDNGDPLNNTTPTVGTASPGGLRELQLSLTEILKGGVLIQRNPOL 180
QY 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240
Db 181 CYQDTILWKDIFHKNNQALTLIDTNRGRACHPCSPMCKGSRGWGSESDCQSLTRTVCA 240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:31:25 ; Search time 16.1672 Seconds
(without alignments)
1772.256 Million cell updates/sec

Title: US-09-821-883-1

Perfect score: 3020

Sequence: 1 MRAAPLLARASLSLGLF.....ENPEYGLDVPAAHHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2410.5	79.8	1255	4	US-09-527-487-2
2	2410.5	79.8	1255	4	US-09-811-115-3
3	2405.5	79.7	1255	1	US-08-467-083-68
4	2405.5	79.7	1255	1	US-08-414-4178-68
5	2405.5	79.7	1255	2	US-08-486-348A-68
6	2405.5	79.7	1255	2	US-08-625-101-2
7	2405.5	79.7	1255	2	US-08-468-545B-68
8	2405.5	79.7	1255	2	US-08-356-786-2
9	2405.5	79.7	1255	3	US-08-466-680B-68
10	2405.5	79.7	1255	4	US-09-354-533-68
11	2385.5	79.0	1255	2	US-08-484-438-8
12	1612	53.4	782	2	US-09-146-283-4
13	1612	53.4	782	3	US-08-579-823A-4
14	1612	53.4	782	3	US-09-344-195-4
15	1607	53.2	419	4	US-09-630-155-2
16	1597	52.9	624	3	US-08-422-108-1
17	1597	52.9	624	4	US-08-422-734-1
18	1183	39.2	580	1	US-08-414-4178-69
19	1183	39.2	580	2	US-08-486-348A-69
20	1183	39.2	580	2	US-08-468-545B-69
21	1183	39.2	580	3	US-08-466-680B-69
22	1183	39.2	580	4	US-09-354-533-69
23	860	28.5	166	4	US-09-648-067A-1
24	721	23.9	644	1	US-08-336-708A-9
25	721	23.9	911	2	US-08-484-438-10
26	721	23.9	1058	2	US-08-484-438-4
27	721	23.9	1210	2	US-08-484-438-7

28	721	23.9	1210	2	US-08-475-035-4	Sequence 4, Appli
29	721	23.9	1308	2	US-08-484-438-2	Sequence 2, Appli
30	704.5	23.3	478	4	US-09-570-454-2	Sequence 2, Appli
31	704.5	23.3	478	4	US-09-867-521-2	Sequence 2, Appli
32	698.5	23.1	1343	6	5183884-4	Patent No. 5183884
33	697	23.1	1342	1	US-07-978-895-4	Sequence 4, Appli
34	697	23.1	1342	2	US-08-484-438-9	Sequence 9, Appli
35	697	23.1	1342	2	US-08-473-119-4	Sequence 4, Appli
36	697	23.1	1342	2	US-08-475-352-4	Sequence 4, Appli
37	697	23.1	1342	4	US-09-170-699-4	Sequence 4, Appli
38	493	16.3	97	1	US-08-421-356-3	Sequence 3, Appli
39	493	16.3	97	4	US-09-046-783-3	Sequence 3, Appli
40	257.5	8.5	1382	2	US-08-737-715-2	Sequence 2, Appli
41	257.5	8.5	1382	4	US-09-457-040B-7	Sequence 7, Appli
42	246	8.1	516	3	US-08-746-559A-4	Sequence 4, Appli
43	246	8.1	1367	2	US-08-249-687C-2	Sequence 2, Appli
44	246	8.1	1367	2	US-08-625-819-2	Sequence 2, Appli
45	246	8.1	1367	3	US-08-746-559A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-527-487-2

; Sequence 2, Application US/09527487

; Patent No. 6528060

; GENERAL INFORMATION:

; APPLICANT: Nicolette, Charles

; TITLE OF INVENTION: HER2 ANTIGENIC PEPTIDES

; FILE REFERENCE: 126881309200

; CURRENT APPLICATION NUMBER: US/09/527,487

; CURRENT FILING DATE: 2000-03-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1255

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-527-487-2

Query Match 79.8%; Score 2410.5; DB 4; Length 1255;

Best Local Similarity 41.2%; Pred. No. 1.1e-183;

Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY	35	LARGAASQVCTGTDMLRLPASPEHLDMLRLHLYQGVVQGNLELYLPTNASLSFLQ	94
DB	16	LPPGAASQVCTGTDMLRLPASPEHLDMLRLHLYQGVVQGNLELYLPTNASLSFLQ	75
QY	95	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPTVTGASPG	154
DB	76	DIQEVQGYVLIHAHQVQVPLQRLRIVRGTLQFEDNYALAVLDNGDPLNNTTPTVTGASPG	135
QY	155	GIREQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIHKNQALTLIDTNRSRACHPC	214
DB	136	GIREQLRLSLTEILKGGVLIQORNPOLCYQDTILWKDIHKNQALTLIDTNRSRACHPC	195
QY	215	SPMKSGSCWGESSEDCQSLRTVTCAGGCARCKGKPLPTDCCHQCAAGCTGPKHSDCLAC	274
DB	196	SPMKSGSCWGESSEDCQSLRTVTCAGGCARCKGKPLPTDCCHQCAAGCTGPKHSDCLAC	255
QY	275	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGS	329
DB	256	LHFNHSGICELHCPALVTYNTDTFESMPNPEGRYTFGASCVTACPNYLSLTDVGS	315
QY	330	-----	329
DB	316	PLHNOEVTAEBDGTQRCCKSPCARVCYGLGMEHLREVRVTSANIQBFAGCKKIFGSLA	375
QY	330	-----	329
DB	376	FLPESFDGDPASNTAPLQEQVETLEETGYLYISANPDLSLPDLVSFQNLQVIRGRI	435

```

QY 330 ----- 329
Db 436 LHNGAYSLTQIGISWGLRSIRLGLSGLALIHNNTHLCFVHTVPWDQLFRNPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGACHQLCARGHCWGPQTQCVCNSQFLRGQECVECRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCINCTHSCVDLDDKGCPCAEQASPLTSIVSAVGLLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYTMRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAGFTVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKIIDEAYVMAGVGSYPVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCLLDHVRENRRGLSGQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRPRELVSE 975
QY 330 ----- 329
Db 976 FSRMARDPQFVVIQNEIDGPASPLDSTFYRSILLEDDMGDLVDAEYLVPOQGFCDP 1035
QY 330 --GAGGMVHRRSSTRSGGDLTLGLPSEEEAPRSPAPSEAGSDVFDGLGMGAA 387
Db 1036 APGAGGMVHRRSSTRSGGDLTLGLPSEEEAPRSPAPSEAGSDVFDGLGMGAA 1095
QY 388 KGLQSLPDPSPLORYSEDPTVPLPSETDGVVAPLTCSPQPEYVQPDVRRPQPSREG 447
Db 1096 KGLQSLPDPSPLORYSEDPTVPLPSETDGVVAPLTCSPQPEYVQPDVRRPQPSREG 1155
QY 448 PLPAARPAGATLERAKTLSPGKNGVVKOVAFGGAVENPEYLTPOGGAAPQHPPPAFSP 507
Db 1156 PLPAARPAGATLERAKTLSPGKNGVVKOVAFGGAVENPEYLTPOGGAAPQHPPPAFSP 1215
QY 508 AFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 546
Db 1216 AFDNLYYWDQDPPERGAPSTFKGTPTAENPEYLGIDVP 1254

```

RESULT 2

```

US-09-811-115-3
; Sequence 3, Application US/09811115
; Patent No. 6632979
; GENERAL INFORMATION:
; APPLICANT: Erickson, Sharon
; APPLICANT: Schwall, Ralph
; APPLICANT: King, Kathleen
; TITLE OF INVENTION: HER-2 TRANSGENIC NON-HUMAN TUMOR MODEL
; FILE REFERENCE: GENENT.034A
; CURRENT APPLICATION NUMBER: US/09/811,115
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,844
; PRIOR FILING DATE: 2000-03-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 3
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-115-3

Query Match      79.8%; Score 2410.5; DB 4; Length 1255;
Best Local Similarity 41.2%; Pred. No. 1.1e-183;
Matches 510; Conservative 0; Mismatches 2; Indels 727; Gaps 1;

QY 35 LARGAASQVCTGDMKRLRSPASBETHLDMRLRHLYQGCVVQGNLELYLTPTNASLSFLQ 94
Db 16 LPPGAASQVCTGDMKRLRSPASBETHLDMRLRHLYQGCVVQGNLELYLTPTNASLSFLQ 75
QY 95 DIOEVQGVVLTAHQVQVPLQRLRIVRGQTOLFDENYALAVLDNGDPLNNTTTPVTGASPG 154
Db 76 DIOEVQGVVLTAHQVQVPLQRLRIVRGQTOLFDENYALAVLDNGDPLNNTTTPVTGASPG 135
QY 155 GLREIQLRSLTEILKGGVLIQRNPQLCYQDTILMKDIFHKNNQALALTLIDNRSRACHPC 214
Db 136 GLREIQLRSLTEILKGGVLIQRNPQLCYQDTILMKDIFHKNNQALALTLIDNRSRACHPC 195
QY 215 SPMCKGSRCSWGESSEDCOSLTRTVACGACGACRCKPLPTDCCHQCACAGCTGPKHSDCLAC 274
Db 196 SPMCKGSRCSWGESSEDCOSLTRTVACGACGACRCKPLPTDCCHQCACAGCTGPKHSDCLAC 255
QY 275 LHFHNSGICELHCPALVTYNTDTFESMPNPRGRTFEGASCVTACPYNYLSTDVGS---- 329
Db 256 LHFHNSGICELHCPALVTYNTDTFESMPNPRGRTFEGASCVTACPYNYLSTDVGSCTLVC 315
QY 330 ----- 329
Db 316 PLHNQEVTAEDTQCEKCKPCARVCYGLGMEHLREVRAVTSANIQEFAGCKKIFGSLA 375
QY 330 ----- 329
Db 376 FLPESEFDGPASNTAPLQEQLVFETLEEITGVLYISAWPDSLPLDSVFNQLQVIRGRI 435
QY 330 ----- 329
Db 436 LHNGAYSLTQIGISWGLRSIRLGLSGLALIHNNTHLCFVHTVPWDQLFRNPHQALLH 495
QY 330 ----- 329
Db 496 TANRPEDECVGEGACHQLCARGHCWGPQTQCVCNSQFLRGQECVECRVLQGLPREYV 555
QY 330 ----- 329
Db 556 NARHCLPCHPECPQNGSVTCFGEADQCVACAHYKDPFCVACPSGVKPDLSYMPIWK 615
QY 330 ----- 329
Db 616 FPDEGACQPCINCTHSCVDLDDKGCPCAEQASPLTSIVSAVGLLVVLGVVFGILI 675
QY 330 ----- 329
Db 676 KRRQOKIRKYTMRLLQETELVEPLTPSGAMPNQAQMRILKETELRKVKVLGSGAGFTVY 735
QY 330 ----- 329
Db 736 KGIWIPDGENVKIPVAIKVLRNTPSKANKIIDEAYVMAGVGSYPVSRLLGICLTSTVQ 795
QY 330 ----- 329
Db 796 LVTQLMPYGCLLDHVRENRRGLSGQDLLNWCQIAKMSYLEDVRLVHRDLAARNVLVKS 855
QY 330 ----- 329
Db 856 PNHVKITDFGLARLLDIDETEHADGGKVPKIKWMALESILRRRFTHQSDVMSYGVTVWEL 915
QY 330 ----- 329
Db 916 MTFGAKPYDGIPAREIPDLLEKGERLPQPICTIDVYIMVKWMIDSECRPRPRELVSE 975

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 12:20:15 ; Search time 61.1253 Seconds
(without alignments)
2565.449 Million cell updates/sec

Title: US-09-821-883-1
Perfect score: 3020
Sequence: 1 MRAAPLLARAAASLSGLF.....ENPYGLDVPAAAHHHH 555

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3020	100.0	555	4	Aae13108 Human HER
2	3005.5	99.5	564	4	Aae13110 Human HER
3	2972	98.4	690	4	Aae13109 Human HER
4	2957.5	97.9	697	4	Aae13111 Human HER
5	2573.5	85.2	919	3	Aab21203 Human HER
6	2573.5	85.2	919	5	Aam51148 Her-2/neu
7	2410.5	79.8	1255	4	Aay92620 Human her
8	2410.5	79.8	1255	4	Aab60167 HER2 tran
9	2410.5	79.8	1255	4	Aae12130 Human tyr
10	2410.5	79.8	1255	5	Aae12130 Human HER
11	2410.5	79.8	1255	5	Aae26349 Human HER
12	2410.5	79.8	1255	5	Aae26366 Human Her
13	2410.5	79.8	1255	5	Aau74545 Human HER
14	2410.5	79.8	1255	6	Abp74708 Breast ca
15	2410.5	79.8	1255	6	Aac38390 Human c-e
16	2410.5	79.8	1255	6	Ada38143 Human erb
17	2410.5	79.8	1255	7	Ada37255 Human Erb
18	2410.5	79.8	1255	7	Adb67621 Human epi
19	2405.5	79.7	1255	2	Aaw01111 HER-2/neu
20	2405.5	79.7	1255	2	Aaw92406 Human HER
21	2405.5	79.7	1255	3	Aay84780 Amino aci
22	2405.5	79.7	1255	3	Aay84780 Human HER
23	2405.5	79.7	1255	4	Ag88267 HER2/neu
24	2405.5	79.7	1255	4	Aab85458 Human HER
25	2405.5	79.7	1255	5	Aae20479 Human HER

26	2405.5	79.7	1255	5	AAU77114	Human Her
27	2405.5	79.7	1255	5	AAAS1143	Human Her
28	2405.5	79.7	1255	5	AAE24067	Human Her
29	2405.5	79.7	1255	6	ABR43687	Human c-e
30	2405.5	79.7	1255	7	ABR82066	Human Her
31	2405.5	79.7	1255	7	ADC09593	Her2/Neu
32	2405.5	79.7	1255	7	ADD25484	Binding d
33	2405.5	79.7	1255	7	ADG63281	Human Pro
34	2405.5	79.7	1255	7	ADE76190	Human HER
35	2398.5	79.4	1253	7	ADC35106	Human bre
36	2389.5	79.1	1433	2	AAR39568	Sequence
37	2334.5	77.3	1223	5	AAU98923	Human bre
38	2129	70.5	920	5	AAAS1152	Mouse Her
39	2129	70.5	926	5	AAAS1153	Mouse Her
40	2092.5	69.3	1200	3	AAAB21208	Human HER
41	1961	64.9	1256	3	AAAB21206	Mouse Her
42	1961	64.9	1256	4	AAAG62860	Amino aci
43	1961	64.9	1256	5	AAAS1151	Mouse Her
44	1950.5	64.6	1260	6	ABU79135	ErbB2 (HE
45	1948	64.5	1256	3	AAAB21199	Rat HER-2

ALIGNMENTS

RESULT 1
AAE13108
ID AAE13108 standard; protein; 555 AA.
XX
AC AAE13108;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human HER500 fusion protein construct.
XX
KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
KW Immunostimulatory component; T-cell mediated immune response; DC;
KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
KW membrane distal intracellular domain; C-terminal tag; human;
KW HER-2 protein; HER500 fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
OS Chimeric.
XX
XX
XX WO200174855-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010515.
XX
XX 30-MAR-2000; 2000US-0193504P.
XX
XX (DEND-) DENDREON CORP.
XX
XX Laus R, Vidovic D, Graddie T;
XX WPI; 2001-662965/76.
XX N-PSDB; AAD21564.
XX
XX An immunostimulatory fusion protein comprising the intracellular domain of HER-2 and an antigen elicits an immune response to the antigen and is useful for the treatment of associated cancer associated.

Claim 7; Page 26; 59pp; English.

The invention relates to immunostimulatory fusion proteins (IFP) and nucleic acid molecules encoding such proteins. The IFPs comprise a polypeptide antigen component and an immunostimulatory component derived from the intracellular domain of HER-2 protein which is effective to elicit a protective dendritic cell (DC)-induced T-cell mediated cellular immune response to the antigen. IFP or superactivated dendritic cells are

CC used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 CC associated with a particularly antigen. The present sequence is HER500
 CC fusion protein construct which comprises human PAP signal sequence,
 CC mature PAP protein, an Ala Arg linker, human HER-2 signal sequence,
 CC mature HER-2 membrane distal extracellular and intracellular domains and
 CC a C-terminal tag
 XX
 XX Sequence 555 AA;
 SQ

Query Match 100.0%; Score 3020; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 6.4e-202;
 Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRAAPLLARAAASLSGLFLFLFWLDRSLAKELARGAASSTVCTGDMKRLRPASPET 60
 Db 1 MRAAPLLARAAASLSGLFLFLFWLDRSLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQCQVQVQGNLELYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRI 120
 Db 61 HLDMLRHLQYQCQVQVQGNLELYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180
 Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180

Qy 181 CYQDTILWKDIFHKNQNLALTIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240
 Db 181 CYQDTILWKDIFHKNQNLALTIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLETDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
 Db 241 GGCARCKGPLETDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGYRTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGGDLTLGLEPSEE 360
 Db 301 MPNPEGYRTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGGDLTLGLEPSEE 360

Qy 361 EAPRSLAPSGAGSDVFDGLGMAAGKLSLTHDPSPLQRYSEDPTVPLPSETDGYV 420
 Db 361 EAPRSLAPSGAGSDVFDGLGMAAGKLSLTHDPSPLQRYSEDPTVPLPSETDGYV 420

Qy 421 APLTCSPOPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNGWVDVAFG 480
 Db 421 APLTCSPOPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNGWVDVAFG 480

Qy 481 GAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQPPRGAPPSTFKGTPTAENPEY 540
 Db 481 GAVENPEYLTPOGGAAPQHPHPPAPSPAFDNLVYWDQPPRGAPPSTFKGTPTAENPEY 540

Qy 541 LGLDVPAAAHHHHHH 555
 Db 541 LGLDVPAAAHHHHHH 555

RESULT 2
 AAEL13110
 ID AAEL13110 standard; protein; 564 AA.
 XX
 AC AAEL13110;
 XX
 XX
 DT 28-JAN-2002 (first entry)
 XX
 DE Human HER500 fusion protein construct comprising OVA-derived octapeptide.
 XX
 KW Immunostimulatory fusion protein; IFP; antigen component; therapy;
 KW immunostimulatory component; T-cell mediated immune response; DC;
 KW dendritic cell; colon cancer; breast carcinoma; ovarian cancer;
 KW PAP protein; Ala Arg linker; membrane distal extracellular domain;
 KW membrane distal intracellular domain; C-terminal tag; human; OVA;
 KW HER-2 protein; ovalbumin-derived octapeptide; HER500 fusion protein.
 XX
 OS Homo sapiens.
 OS Unidentified.

OS Chimeric.
 XX WO200174855-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US010515.
 XX 30-MAR-2000; 2000US-0193504P.
 XX (DEND-) DENDREON CORP.
 XX Laus R, Vidovic D, Graddis T;
 XX WPI; 2001-622965/76.
 XX N-PSDB; AAD21566.
 XX An immunostimulatory fusion protein comprising the intracellular domain
 XX of HER-2 and an antigen elicits an immune response to the antigen and is
 XX useful for the treatment of associated cancer associated.
 XX Claim 7; Page 26; 59pp; English.
 XX The invention relates to immunostimulatory fusion proteins (IFP) and
 XX nucleic acid molecules encoding such proteins. The IFPs comprise a
 XX polypeptide antigen component and an immunostimulatory component derived
 XX from the intracellular domain of HER-2 protein which is effective to
 XX elicit a protective dendritic cell (DC)-induced T-cell mediated cellular
 XX immune response to the antigen. IFP or superactivated dendritic cells are
 XX used to treat cancer e.g. breast carcinoma, ovarian and colon cancer
 XX associated with a particularly antigen. The present sequence is HER500
 XX fusion protein construct which comprises human PAP signal sequence,
 XX mature PAP protein, an Ala Arg linker, human HER-2 signal sequence,
 XX mature HER-2 membrane distal extracellular domain, an Ala linker, an
 XX ovalbumin (OVA)-derived immunodominant octapeptide, HER-2 membrane distal
 XX intracellular domain and a C-terminal tag
 XX Sequence 564 AA;
 SQ

Query Match 99.5%; Score 3005.5; DB 4; Length 564;
 Best Local Similarity 99.4%; Pred. No. 6.7e-201;
 Matches 555; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 MRAAPLLARAAASLSGLFLFLFWLDRSLAKELARGAASSTVCTGDMKRLRPASPET 60
 Db 1 MRAAPLLARAAASLSGLFLFLFWLDRSLAKELARGAASSTVCTGDMKRLRPASPET 60

Qy 61 HLDMLRHLQYQCQVQVQGNLELYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRI 120
 Db 61 HLDMLRHLQYQCQVQVQGNLELYLPTNASLSFLQDIQEVQGYVLIHNNQVQVPLQRLRI 120

Qy 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180
 Db 121 VRGTQLFEDNYALAVLDNGDPLNNTPTVTGASPGGLRELQRLSLTEILKGGVLIQRNPOL 180

Qy 181 CYQDTILWKDIFHKNQNLALTIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240
 Db 181 CYQDTILWKDIFHKNQNLALTIDTNRSRACHPCSPMKSGRCWGESSEDCQSLTRTVCA 240

Qy 241 GGCARCKGPLETDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300
 Db 241 GGCARCKGPLETDCHEQCAAGCTGPKHSDCLACLFHNSGICELHCPALVYNTDTPES 300

Qy 301 MPNPEGYRTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGGDL 351
 Db 301 MPNPEGYRTFGASCVTACPNYLSLTDVSGAGGMVHHRSSSTRSGGGDL 360

Qy 352 TLGLEPSEERAPRSLAPSEAGSDVFDGLGMAAGKLSLTHDPSPLQRYSEDPTVP 411
 Db 361 TLGLEPSEERAPRSLAPSEAGSDVFDGLGMAAGKLSLTHDPSPLQRYSEDPTVP 420

Qy 412 LPSETDGYVAPLTCSPQPEYVNDVPRPQPPSPREGPLPAARAGATLERAKTSLSPGKNG 471
 XX